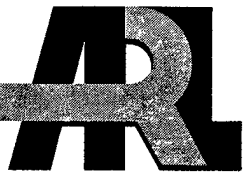
ARMY RESEARCH LABORATORY



Spread-Spectrum Code-Division Multiple Access

Don Torrieri

ARL-TR-2721

July 2002

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Adelphi, MD 20783-1197

ARL-TR-2721 **July 2002**

Spread-Spectrum Code-Division Multiple Access

Don Torrieri

Computational and Information Sciences Directorate

Preface

This report is intended to provide a concise but lucid explanation and derivation of the fundamentals of spread-spectrum code-division multiple access (CDMA). The level of presentation is suitable for those with a solid background in the theory of digital communications. Throughout the report, there are many streamlined derivations, new derivations, and simplifications of the classical theory. Comprehensive mathematical details, some of them difficult to find elsewhere, are assembled in the appendices. Sections 1 and 2 contain brief overviews of fading and diversity that provide the basic theory accessed by the remainder of the report. Maximal-ratio combining, equal-gain combining, and selection diversity are distinguished and compared. The rake receiver and the impact of error-correcting codes on fading communications are examined. Direct-sequence CDMA is considered in Section 3. Bit error probabilities for a number of direct-sequence systems are derived. Power-control issues in cellular networks are emphasized and new results are presented. Diversity aspects of multicarrier and single-carrier systems are compared. Section 4 treats frequency-hopping CDMA for both peer-to-peer and cellular systems. The advantages of frequency hopping in network applications and the effects of spatial diversity, spectral splatter, and the number of equivalent frequency channels are explained.

Contents

Figures

Tables

1. Fading for Wireless Communications

1.1 Path Loss and Shadowing

Free-space propagation losses of electromagnetic waves vary inversely with the square of the distance between a transmitter and a receiver. For terrestrial propagation, when one signal traversing a direct path and another perfectly reflected from a plane interfere at a receiver, it is found that the propagation loss varies as the inverse of the fourth-power of the distance. Thus, it is natural to seek a power-law variation for the average received power in a specified geographic area as a function of distance. For terrestrial wireless communications, measurements averaged over many different positions of a transmitter and a receiver in a specified geographic area confirm that the average received power, which is called the *area-mean power*, does tend to vary inversely as a power of the transmitter-receiver distance r. It is found that the area-mean power is approximately given by

$$p_a = p_0 \left(\frac{r}{R_0} \right)^{-\beta} \tag{1-1}$$

where p_0 is the average received power when the distance is $r = R_0$, and β is the *attenuation power law*. The parameters p_0 and β are functions of the carrier frequency, antenna heights, terrain characteristics, vegetation, and various characteristics of the propagation medium. Typically, the parameters vary with distance, but are constant within a range of distances. A typical value of the attenuation power law for urban areas and microwave frequencies is $\beta = 4$. The power law increases with the carrier frequency.

For a specific propagation path, the received *local-mean power* departs from the area-mean power due to *shadowing*, which is the effect of diffractions and propagation conditions that are path-dependent. Each diffraction due to obstructing terrain and each reflection from an obstacle causes the signal power to be multiplied by an attenuation factor. Thus, the received signal power is often the product of many factors, and hence the logarithm of the signal power is the sum of many factors. If each factor is modeled as a uniformly bounded, independent random variable that varies from path to path, then the central-limit theorem implies that the logarithm of the received signal power has an approximately normal distribution if the number of attenuation factors is large enough. Extensive empirical data confirms that the received local-mean power after transmission over a randomly selected propagation path with a fixed distance is approximately lognormally distributed. Thus, the shadowing model specifies that the local-mean power has the form

$$p_l = p_a 10^{\xi/10} \tag{1-2}$$

where the *shadowing factor* ξ is a zero-mean random variable with a normal distribution. The standard deviation of ξ is denoted by σ_s, which is expressed in decibels. From (1-1) and (1-2), it follows that the probability distribution function of the normalized local-mean power, p_l/p_0, is

$$F(x) = 1 - \frac{1}{2}\text{erfc}\left\{\frac{a}{\sigma_s}\ln\left[x\left(\frac{r}{R_0}\right)^{\beta}\right]\right\} \tag{1-3}$$

where erfc{ } denotes the complementary error function, ln[ ] denotes the natural logarithm, and $a = (10\log_{10}e)/\sqrt{2}$. The standard deviation σ_s increases with carrier frequency and terrain irregularity and often exceeds 10 dB for terrestrial communications. The value of the shadowing factor for a propagation path is usually strongly correlated with its value for a nearby propagation path. For mobile communications, the typical time interval during which the shadowing factor is nearly constant is a second or more.

Fading, which is endemic in mobile, long-distance, high-frequency, and other communication channels, causes power fluctuations about the local-mean power. Fading occurs at much faster rate than shadowing. During an observation interval in which the shadowing factor is nearly constant, the received signal power may be expressed as the product

$$p_r = p_a 10^{\xi/10}\alpha^2(t) \tag{1-4}$$

where the factor $\alpha^2(t)$ is due to the fading. Since ξ is fixed, the local-mean power is

$$p_l = E[p_r] = p_a 10^{\xi/10}E[\alpha^2(t)] \tag{1-5}$$

where E[ ] denotes the expected value.

1.2 Fading

A signal experiences *fading* when the interaction of multipath components and time- or frequency-varying channel conditions cause significant fluctuations in its amplitude at a receiver. *Multipath components* of a signal are generated by inhomogeneities in the propagation medium or reflections from obstacles. These components travel along different paths before being recombined at the receiver. Because of the different time-varying delays and attenuations encountered by the multipath components, the recombined signal is a distorted version of the original transmitted signal. Fading may be classified as time-selective, frequency-selective, or both. *Time-selective fading* is fading caused by the movement of the transmitter or receiver or by changes in the propagation medium. *Frequency-selective fading* is fading caused by the different delays of the multipath components, which may affect certain frequencies more than others. The following concise development of fading theory emphasizes basic physical mechanisms. More general theoretical frameworks [1], [2], [3], are available.

A bandpass transmitted signal can be expressed as

$$s_t(t) = \text{Re}[m(t)\exp(j2\pi f_c t)] \tag{1-6}$$

where $m(t)$ denotes its complex envelope, f_c denotes its carrier frequency, and Re[ ] denotes the real part. Transmission over a time-varying multipath channel of $N(t)$ paths produces a received bandpass signal that consists of the sum of $N(t)$ waveforms. The ith waveform is the transmitted signal delayed by time $\tau_i(t)$, attenuated by a factor $a_i(t)$ that depends on the path loss and shadowing, and shifted in frequency by the amount $f_{di}(t)$ due to the Doppler effect. Assuming that $f_{di}(t)$ is constant during the path delays, the received signal may be expressed as

$$s_r(t) = \text{Re}[r_1(t)\exp(j2\pi f_c t)] \tag{1-7}$$

where the received complex envelope is

$$r_1(t) = \sum_{i=1}^{N(t)} a_i(t)\exp[j\phi_i(t)]m[t-\tau_i(t)] \tag{1-8}$$

and its phase is

$$\phi_i(t) = -2\pi f_c \tau_i(t) + 2\pi f_{di}(t)\,[t-\tau_i(t)] \tag{1-9}$$

The Doppler shift arises because of the relative motion between the transmitter and the receiver. In Figure 1(a), the receiver is moving at speed $v(t)$ and the angle between the velocity vector and the propagation direction of an electromagnetic wave is $\psi_i(t)$. Thus, the received frequency is increased by the Doppler shift

$$f_{di}(t) = f_c\frac{v(t)}{c}\cos\psi_i(t) \tag{1-10}$$

where c is the speed of an electromagnetic wave. In Figure 1(b), the transmitter is moving at speed $v(t)$ and there is a reflecting surface that changes the arrival angle of the electromagnetic wave at the receiver. If $\psi_i(t)$ represents the angle between the velocity vector and the initial direction of the electromagnetic wave, then (1-10) again gives the Doppler shift.

1.3 Time-Selective Fading

To analyze time-selective fading, it is assumed that $N(t) = N$ for the time interval of interest and that the differences in the time delays along the various paths are small compared with the inverse of the signal bandwidth. Therefore, the received multipath components overlap in time and are called *unresolvable multipath components*. If the time origin is chosen to coincide with the arrival time of the first multipath component at a receiver, then the received complex envelope of (1-8) may be expressed as

$$r_1(t) \approx m(t)r(t) \tag{1-11}$$

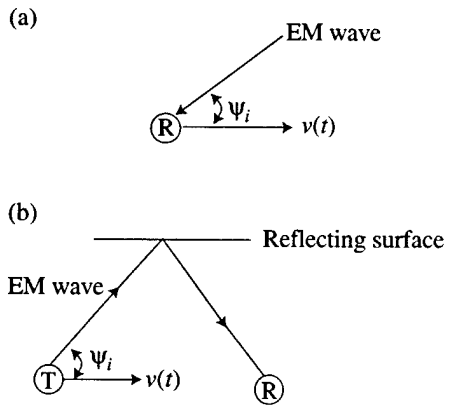


Figure 1. Examples of the Doppler effect: (a) receiver motion and (b) transmitter motion and reflecting surface.

where the *equivalent lowpass* or *equivalent baseband channel response* is

$$r(t) = \sum_{i=1}^{N} a_i(t) \exp\left[j\phi_i(t)\right] \tag{1-12}$$

The fluctuations in this factor cause signal fading at the receiver and increase the bandwidth of the received signal. If the transmitted signal is an unmodulated tone, then $m(t) = 1$ and (1-12) represents the complex envelope of the received signal.

The channel response can be decomposed as

$$r(t) = r_c(t) + jr_s(t) \tag{1-13}$$

where $j = \sqrt{-1}$ and

$$r_c(t) = \sum_{i=1}^{N} a_i(t) \cos[\phi_i(t)] \,, \qquad r_s(t) = \sum_{i=1}^{N} a_i(t) \sin[\phi_i(t)] \tag{1-14}$$

If the range of the delay values exceeds $1/f_c$, then the sensitivity of $\phi_i(t)$ to small variations in the delay $\tau_i(t)$ makes it plausible to model the phases $\phi_i(t), i = 1, 2, \ldots, N$, as random variables that are independent of each other and the $\{a_i(t)\}$ and are uniformly distributed over $[0, 2\pi)$ at a specific time t. Therefore,

$$E[r_c(t)] = E[r_s(t)] = 0 \tag{1-15}$$

If the amplitude factors $a_i(t), i = 1, 2, \ldots, N$, are either identically distributed or uniformly bounded independent random variables at time t, then according to the central-limit theorem, the probability distributions of both $r_c(t)$ and $r_s(t)$ approach Gaussian distributions as N increases. Thus, if N is sufficiently large, then $r(t)$ at a specific time is well modeled as a complex Gaussian random variable. Since the phases are independent and uniformly distributed, it follows that

$$E[r_c(t)r_s(t)] = 0 \tag{1-16}$$

$$E[r_c^2(t)] = E[r_s^2(t)] = \sigma_r^2(t) \tag{1-17}$$

where we define

$$\sigma_r^2(t) = \frac{1}{2}\sum_{i=1}^{N} E[a_i^2(t)] \tag{1-18}$$

This equation indicates that $\sigma_r^2(t)$ is equal to the sum of the local-mean powers of the multipath components. Equations (1-15) to (1-17) imply that $r_c(t)$ and $r_s(t)$ are independent, identically distributed, zero-mean Gaussian random variables.

Let $\alpha(t) = |r(t)|$ denote the envelope, and $\theta(t) = \tan^{-1}[r_s(t)/r_c(t)]$ the phase of $r(t)$ at a specific time t. Then

$$r(t) = \alpha(t)e^{j\theta(t)} \tag{1-19}$$

As shown in Appendix A-4, since $r_c(t)$ and $r_s(t)$ are Gaussian and $\alpha^2(t) = r_c^2(t) + r_s^2(t)$, $\theta(t)$ has a uniform distribution over [0, 2π), and $\alpha(t)$ has the *Rayleigh* probability density function:

$$f_\alpha(r) = \frac{r}{\sigma_r^2}\exp\left(-\frac{r^2}{2\sigma_r^2}\right)u(r) \tag{1-20}$$

where the time-dependence has been suppressed for convenience, and $u(r) = 1$, $r \geq 0$, and $u(r) = 0$, $r < 0$. From (1-20) or directly from (1-13) and (1-17), it follows that

$$E[\alpha^2(t)] = 2\sigma_r^2(t) \tag{1-21}$$

The substitution of (1-19) and (1-11) into (1-7) gives

$$\begin{aligned} s_r(t) &= \mathrm{Re}[\alpha(t)\, m(t)\exp(j2\pi f_c t + j\theta(t))] \\ &= \alpha(t)\, A_m(t)\cos[2\pi f_c t + \phi_m(t) + \theta(t)] \end{aligned} \tag{1-22}$$

where $A_m(t)$ is the amplitude and $\phi_m(t)$ the phase of $m(t)$. Equations (1-21) and (1-22) indicate that the instantaneous local-mean power is $p_l = \sigma_r^2(t)A_m^2(t)$.

When a line-of-sight exists between a transmitter and a receiver, one of the received multipath components may be much stronger than the others. This strong component is called the *specular component* and the other unresolvable components are called *diffuse* or *scattered components*. As a result, the multiplicative channel response of (1-12) becomes

$$r(t) = a_0(t)\exp[j\phi_0(t)] + \sum_{i=1}^{N} a_i(t)\exp[j\phi_i(t)] \tag{1-23}$$

where the summation term is due to the diffuse components, and the first term is due to the specular component. If N is sufficiently large, then at time t the summation term is well-approximated by a zero-mean, complex

Gaussian random variable. Thus, $r(t)$ at a specific time is a complex Gaussian random variable with a nonzero mean equal to the deterministic first term, and (1-13) implies that

$$E[r_c(t)] = a_0(t)\cos[\phi_0(t)]\,, \qquad E[r_s(t)] = a_0(t)\sin[\phi_0(t)] \tag{1-24}$$

As shown in Appendix A-3, since $r_c(t)$ and $r_s(t)$ are Gaussian and $\alpha^2(t) = r_c^2(t) + r_s^2(t)$, the envelope $\alpha(t) = |r(t)|$ has the *Rice* probability density function:

$$f_\alpha(r) = \frac{r}{\sigma_r^2}\exp\left\{-\frac{r^2+a_0^2}{2\sigma_r^2}\right\} I_0\left(\frac{a_0 r}{\sigma_r^2}\right) u(r) \tag{1-25}$$

where $I_0(\ )$ is the modified Bessel function of the first kind and order zero, and the time-dependence is suppressed for convenience. From (1-25) or directly from (1-18) and (1-23), it follows that the average envelope power is

$$\Omega = E[\alpha^2(t)] = a_0^2(t) + 2\sigma_r^2(t) \tag{1-26}$$

The type of fading modeled by (1-23) and (1-25) is called *Ricean fading*. At a specific time, the *Rice factor* is defined as

$$\kappa = \frac{a_0^2}{2\sigma_r^2} \tag{1-27}$$

which is the ratio of the specular power to the diffuse power. In terms of κ and Ω, the Rice density is

$$f_\alpha(r) = \frac{2(\kappa+1)}{\Omega} r \exp\left\{-\kappa - \frac{(\kappa+1)r^2}{\Omega}\right\} I_0\left(\sqrt{\frac{\kappa(\kappa+1)}{\Omega}}2r\right) u(r) \tag{1-28}$$

When $\kappa = 0$, Ricean fading is the same as Rayleigh fading. When $\kappa = \infty$, there is no fading.

A more flexible fading model is created by introducing a new parameter m; the *Nakagami-m* probability density function for the envelope $\alpha(t)$ is

$$f_\alpha(r) = \frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^m r^{2m-1}\exp\left(-\frac{m}{\Omega}r^2\right)u(r)\,,\ m \geq \frac{1}{2} \tag{1-29}$$

where the gamma function $\Gamma(m)$ is defined by (A-12). When $m = 1$, the Nakagami density becomes the Rayleigh density, and when $m \to \infty$, there is no fading. When $m = 1/2$, the Nakagami density becomes the one-sided Gaussian density. A measure of the severity of the fading is $\text{var}(\alpha^2)/(E[\alpha^2])^2$. Equating this ratio for the Rice and Nakagami densities, it is found that when $m \geq 1$, the Nakagami density closely approximates a Rice density with

$$\kappa = \frac{\sqrt{m^2-m}}{m-\sqrt{m^2-m}}\,, \quad m \geq 1 \tag{1-30}$$

Since the Nakagami-m model essentially incorporates the Rayleigh and Rice models as special cases and provides for many other possibilities, it is not surprising that this model often fits well with empirical data. Integrating over (1-29), changing the integration variable, and using (A-12), we obtain

$$E[\alpha^n] = \frac{\Gamma(m+\frac{n}{2})}{\Gamma(m)}\left(\frac{\Omega}{m}\right)^{n/2} \tag{1-31}$$

Consider a time interval small enough that $N(t) = N, v(t) = v$, and $\psi_i(t) = \psi_i$ are approximately constants and $a_i(t) = a_i$ and $\tau_i(t) = \tau_i$ are random variables. Then (1-9) and (1-10) yield

$$\phi_i(t+\tau) - \phi_i(t) = 2\pi f_d \tau \cos\psi_i \tag{1-32}$$

where $f_d = f_0\, v/c$ is the maximum Doppler shift and τ is a time shift. The autocorrelation of a wide-sense-stationary complex process $r(t)$ is defined as

$$A_r(\tau) = \frac{1}{2}E[r^*(t)r(t+\tau)] \tag{1-33}$$

where the asterisk denotes the complex conjugate. The variation of the autocorrelation of the equivalent baseband channel response defined by (1-12) provides a measure of the changing channel characteristics. To interpret the meaning of (1-33), we substitute (1-13) and decompose the autocorrelation as

$$\mathrm{Re}\{A_r(\tau)\} = \frac{1}{2}\{E[r_c(t)r_c(t+\tau)] + E[r_s(t)r_s(t+\tau)]\} \tag{1-34}$$

$$\mathrm{Im}\{A_r(\tau)\} = \frac{1}{2}\{E[r_c(t)r_s(t+\tau)] - E[r_s(t)r_c(t+\tau)]\} \tag{1-35}$$

Thus, the real part of this autocorrelation is the average of the autocorrelations of the real and imaginary parts of $r(t)$; the imaginary part is proportional to the difference between two crosscorrelations of the real and imaginary parts of $r(t)$. Substituting (1-12) into (1-33), using the independence and uniform distribution of each ϕ_i and the independence of a_i and ϕ_i, and then substituting (1-32), we obtain

$$A_r(\tau) = \frac{1}{2}\sum_{i=1}^{N} E[a_i^2]\exp(j2\pi f_d\,\tau\cos\psi_i) \tag{1-36}$$

If all the received multipath components have approximately the same power and the receive antenna is omnidirectional, then (1-18) implies that $E[a_i^2] \approx 2\,\sigma_r^2/N, i = 1, 2, \ldots, N$, and (1-36) becomes

$$A_r(\tau) = \frac{\sigma_r^2}{N}\sum_{i=1}^{N}\exp(j2\pi f_d\,\tau\cos\psi_i) \tag{1-37}$$

A communication system such as a mobile that receives a signal from an elevated base station may be surrounded by many scattering objects. An *isotropic scattering* model assumes that multipath components of comparable power are reflected from many different scattering objects and hence arrive from many different directions. For two-dimensional isotropic scattering, N is large, and the $\{\psi_i\}$ lie in a plane and have values that are uniformly distributed over $[0, 2\pi)$. Therefore, the summation in (1-37) can be approximated by an integral; that is,

$$A_r(\tau) \approx \frac{\sigma_r^2}{2\pi} \int_0^{2\pi} \exp(j2\pi f_d \tau \cos\psi) d\psi \tag{1-38}$$

An evaluation of this integral gives the *autocorrelation of the channel response for two-dimensional isotropic scattering*:

$$A_r(\tau) = \sigma_r^2 J_0(2\pi f_d \tau) \tag{1-39}$$

where $J_0(\,)$ is the Bessel function of the first kind and order zero.

The normalized autocorrelation $A_r(\tau)/A_r(0)$, which is a real-valued function of $f_d\tau$, is plotted in Figure 2. It is observed that its magnitude is less than 0.3 when $f_d\tau > 1$. This observation leads to the definition of the *coherence time* or *correlation time* of the channel as

$$T_c = \frac{1}{f_d} \tag{1-40}$$

where f_d is the maximum Doppler shift or *Doppler spread*. The coherence time is a measure of the time separation between signal samples sufficient for the samples to be largely decorrelated. If the coherence time is much longer than the duration of a channel symbol, then the fading is relatively constant over a symbol and is called *slow fading*. Conversely, if the coherence time is on the order of the duration of a channel symbol or less, then the fading is called *fast fading*.

The power spectral density of a complex process is defined as the Fourier transform of its autocorrelation. From (1-39) and tabulated Fourier transforms, we obtain the *Doppler power spectrum for two-dimensional isotropic scattering*:

$$S_r(f) = \begin{cases} \dfrac{\sigma_r^2}{\pi\sqrt{f_d^2 - f^2}} & , \ |f| < f_d \\ 0 & , \text{ otherwise} \end{cases} \tag{1-41}$$

The normalized Doppler spectrum $S_r(f)/S_r(0)$, which is plotted in Figure 3 versus f/f_d, is bandlimited by the Doppler spread f_d and tends to infinity as f approaches $\pm f_d$. The Doppler spectrum is the superposition of contributions from multipath components, each of which experiences a different Doppler shift upper bounded by f_d.

Figure 2.
Autocorrelation of $r(t)$
for isotropic scattering.

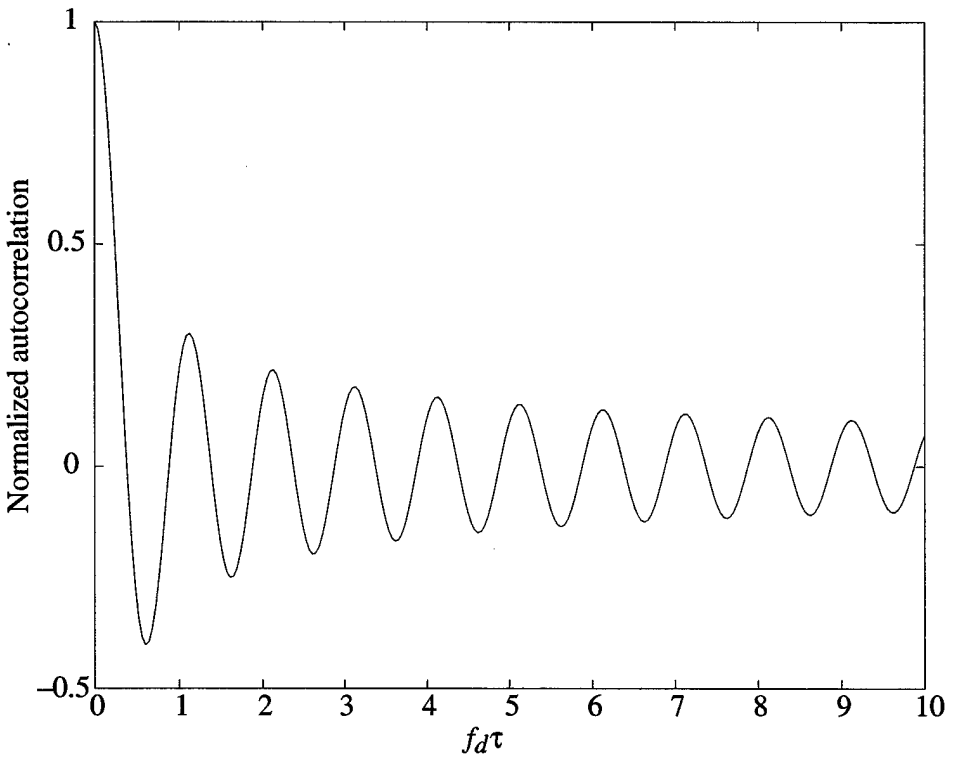
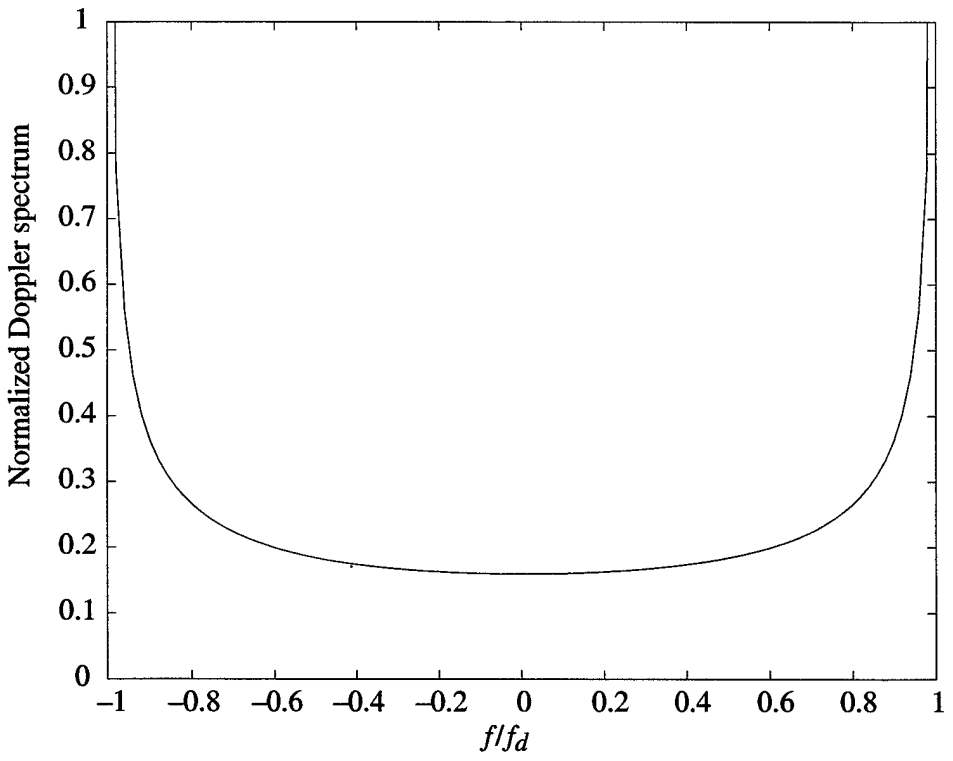


Figure 3. Doppler
spectrum for isotropic
scattering.



The received signal power spectrum may be calculated from (1-7), (1-11), and (1-41). For an unmodulated carrier, $m(t) = 1$ and the received signal power spectrum is

$$S_{\text{rec}}(f) = \frac{1}{2}S_r(f - f_c) + \frac{1}{2}S_r(f + f_c) \tag{1-42}$$

In general, when the scattering is not isotropic, the imaginary part of the autocorrelation $A_r(\tau)$ is nonzero, and the real part decreases much more slowly and less smoothly with increasing τ than (1-39). Both the real and imaginary parts often exhibit minor peaks for time shifts exceeding $1/f_d$. Thus, the coherence time provides only a rough characterization of the channel behavior.

A received signal from one source can often be decomposed into the sum of signals reflected from several clusters of scatterers. The *equivalent complex-valued baseband impulse response* of the channel $h(\tau; t)$ is the response at time t due to an impulse applied τ seconds earlier. For most practical applications, the *wide-sense stationary, uncorrelated scattering* model is reasonably accurate. In this model, the impulse response can be expressed as

$$h(\tau; t) = \sum_{i=1}^{L_c(t)} r_i(t)\,\delta(\tau - \tau_i(t)) \tag{1-43}$$

where $L_c(t)$ is the number of clusters, $\tau_i(t)$ is the delay associated with the ith cluster, $\delta(\,)$ is the Dirac delta function, and $r_i(t)$ is a complex, wide-sense stationary, Gaussian random process with zero mean for Rayleigh fading and a nonzero mean for Ricean fading. Because of the uncorrelated scattering that generates each term in (1-43), the process $r_i(t)$ is uncorrelated with $r_j(t), j \neq i$. Diffuse fading paths can be modeled by replacing one or more terms in (1-43) by integrals over density functions. Both more general and more specialized, but much less often used, models have been proposed by many authors.

1.4 Spatial Diversity and Fading

To obtain spatial diversity in a fading environment, the antennas in an array must be separated enough that there is little correlation between signal replicas or copies at the antennas. To determine what separation is needed, consider the reception of a signal at two antennas separated by a distance D, as illustrated in Figure 4. If the signal arrives as an electromagnetic plane wave, then the signal copy at antenna 1 relative to antenna 2 is delayed by $D \sin\theta/c$, where θ is the arrival angle of the plane wave relative to a line perpendicular to the line joining the two antennas. Let $\phi_{ki}(t)$ denote the phase of the complex envelope of multipath component i at antenna k. Consider a time interval small enough that $\phi_{ki}(t) = \phi_{ki}, N(t) = N, a_i(t) = a_i$, and each multipath component arrives from a fixed angle. Thus, if multipath component i of a narrowband signal arrives as a plane wave at angle ψ_i, then

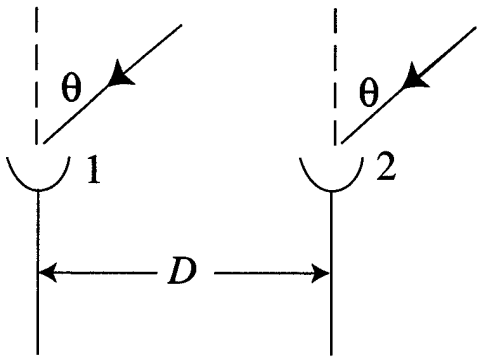


Figure 4. Two antennas receiving plane wave that results in a signal copy at each antenna.

the phase ϕ_{2i} of the complex envelope of the component copy at antenna 2 is related to the phase ϕ_{1i} at antenna 1 by

$$\phi_{2i} = \phi_{1i} + 2\pi \frac{D}{\lambda} \sin \psi_i \tag{1-44}$$

where $\lambda = c/f_c$ is the wavelength of the signal. If the multipath component propagates over a distance much larger than the separation between the two antennas, then it is reasonable to assume that the attenuation a_i is identical at the two antennas. If the range of the delay values exceeds $1/f_c$, then the sensitivity of the phases to small delay variations makes it plausible that for $k = 1, 2$ the phases ϕ_{ki}, $i = 1, 2, \ldots, N$, are well modeled as independent random variables that are uniformly distributed over $[0, 2\pi)$. From (1-12), the complex envelope r_k of the signal copy at antenna k when the signal is a tone is

$$r_k = \sum_{i=1}^{N} a_i \exp(j\phi_{ki}), \quad k = 1, 2 \tag{1-45}$$

The cross-correlation between r_1 and r_2 is defined as

$$C_{12}(D) = \frac{1}{2} E[r_1^* r_2] \tag{1-46}$$

Substituting (1-45) into (1-46), using the independence of each a_i and ϕ_{ki}, the independence of ϕ_{ki} and ϕ_{kl}, $i \neq l$, and the uniform distribution of each ϕ_{ki}, and then substituting (1-44), we obtain

$$C_{12}(D) = \frac{1}{2} \sum_{i=1}^{N} E[a_i^2] \exp(j2\pi D \sin \psi_i / \lambda) \tag{1-47}$$

This equation for the cross-correlation as a function of spatial separation clearly resembles (1-36) for the autocorrelation as a function of time delay. If all the multipath components have approximately the same power so that $E[a_i^2] \approx 2\sigma_r^2/N$, $i = 1, 2, \ldots, N$, then

$$C_{12}(D) = \frac{\sigma_r^2}{N} \sum_{i=1}^{N} \exp(j2\pi D \sin \psi_i / \lambda) \tag{1-48}$$

Applying the two-dimensional isotropic scattering model, a derivation similar to that of (1-39) gives the real-valued crosscorrelation

$$C_{12}(D) = \sigma_r^2 J_0(2\pi D/\lambda) \tag{1-49}$$

This model indicates that an antenna separation of $D \geq \lambda/2$ ensures that the normalized crosscorrelation $C_{12}(D)/C_{12}(0)$ is less than 0.3. A plot of the normalized crosscorrelation is obtained from Figure 2 if the abscissa is interpreted as D/λ. When the scattering is not isotropic or the number of scattering objects producing multipath components is small, then the real and imaginary parts of the crosscorrelation decrease much more slowing with D/λ. For example, Figure 5 shows the real and imaginary parts of the normalized crosscorrelation when the$\{\psi_i\}$ are a nearly continuous band of angles between $7\pi/32$ and $9\pi/32$ radians so that (1-48) can be approximated by an integral. Figure 6 depicts the real and imaginary parts of the normalized crosscorrelation when $N = 9$ and the $\{\psi_i\}$ are uniformly spaced throughout the first two quadrants: $\psi_i = (i-1)\pi/8$, $i = 1, 2, \ldots, 9$. In the example of Figure 5, an antenna separation of at least 5λ is necessary to ensure approximate decorrelation of the signal copies and obtain spatial diversity. In the example of Figure 6, not even a separation of 10λ is adequate to ensure approximate decorrelation.

1.5 Frequency-Selective Fading

Frequency-selective fading occurs because multipath components combine destructively at some frequencies, but constructively at others. The different path delays cause *dispersion* of a received pulse in time and cause intersymbol interference between successive symbols. When a multipath channel introduces neither time-variations nor Doppler shifts, (1-8) and (1-9) indicate that the received complex envelope is

$$r_1(t) = \sum_{i=1}^{L_s} a_i \exp(-j2\pi f_c \tau_i) m(t-\tau_i) \tag{1-50}$$

The number of multipath components L_s includes only those components with power that is a significant fraction, perhaps 0.05 or more, of the power of the dominant component. The multipath *delay spread* T_d is defined as the maximum delay of a significant multipath component relative to the minimum delay of a component; that is,

$$T_d = \max_i \tau_i - \min_i \tau_i \ , \quad i = 1, 2, \ldots, N \tag{1-51}$$

If the duration of a received symbol T_s is much larger than T_d, then the multipath components are usually unresolvable, $m(t-\tau_i) \approx m(t-\tau_1)$, $i = 1, 2, \ldots, L_s$, and hence $r_1(t)$ is proportional to $m(t-\tau_1)$. Since all frequency components of the received signal fade nearly simultaneously, this type of fading is called *frequency-nonselective* or *flat fading* and occurs if $T_s >> T_d$. In contrast, a signal is said to experience *frequency-selective fading* if $T_s < T_d$ because then the time variation or fading of the spectral components of

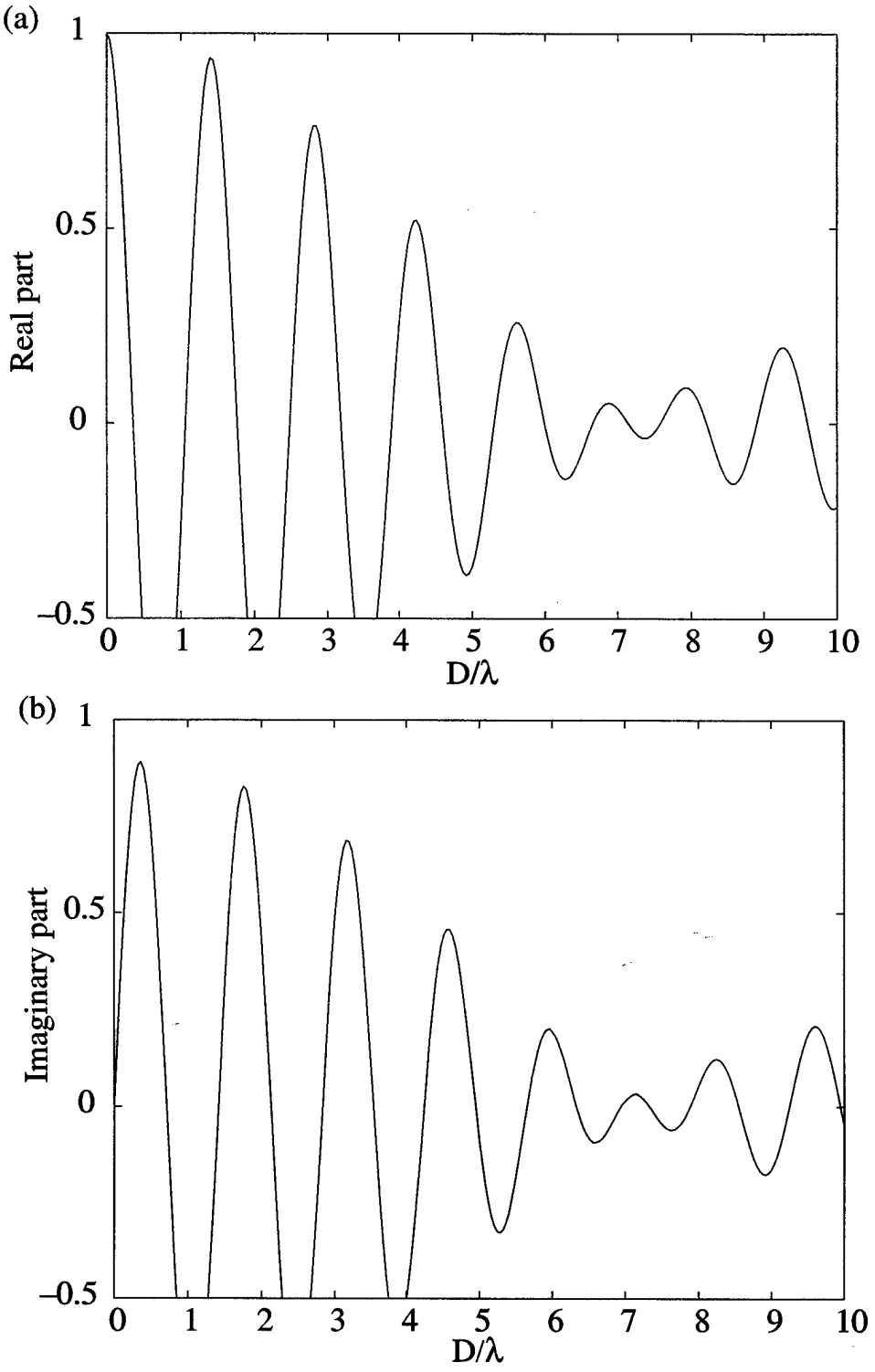


Figure 5. Normalized crosscorrelation for multipath components arriving from band of angles between $7\pi/32$ and $9\pi/32$ radians: (a) real part and (b) imaginary part.

$m(t)$ may be different. The large delay spread may cause intersymbol interference, which is accommodated by equalization in the receiver. However, if the time delays are sufficiently different among the multipath components that they are *resolvable* at the demodulator or matched-filter output,

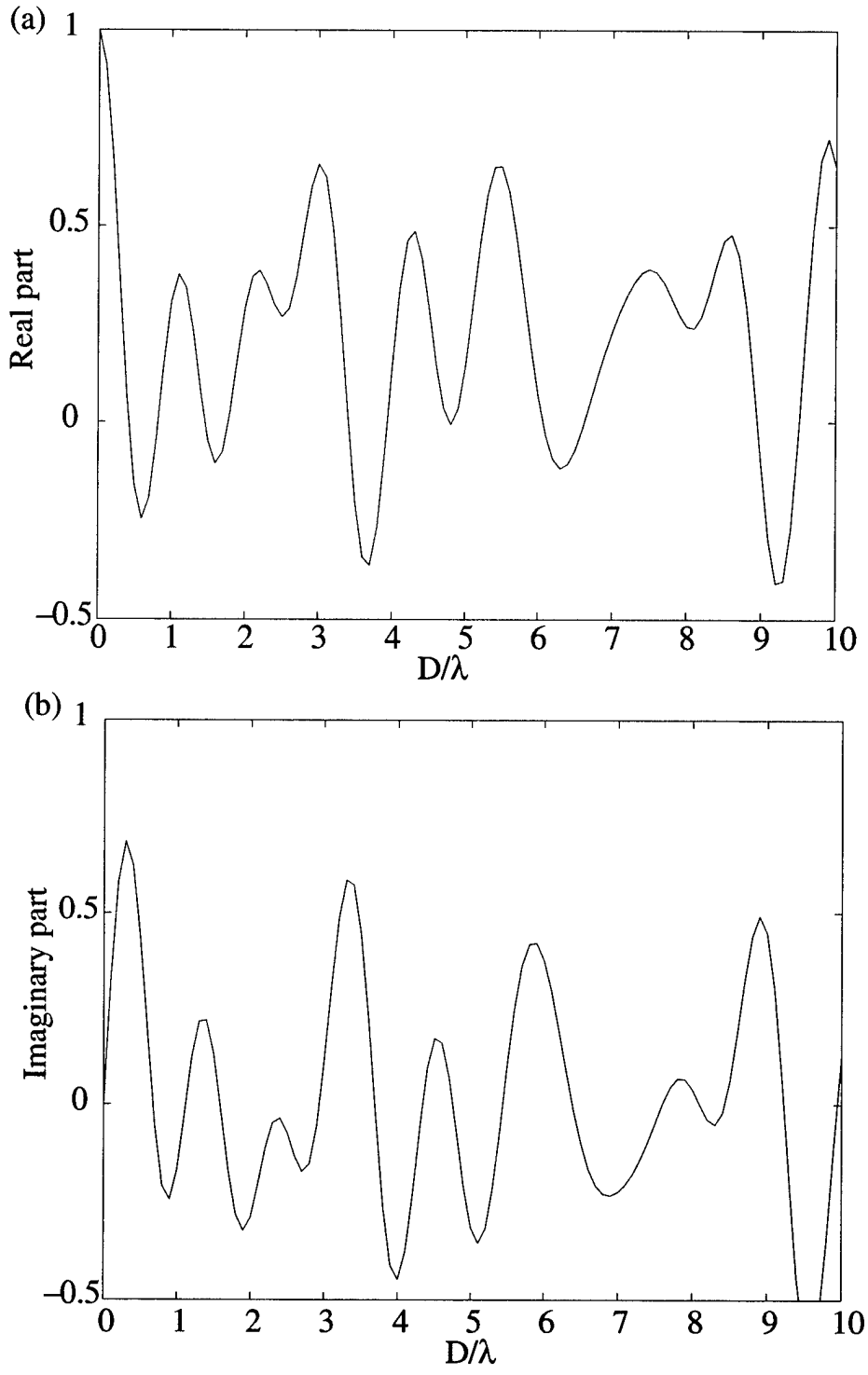


Figure 6. Normalized crosscorrelation for N = 9 multipath components arriving from uniformly spaced angles in the first two quadrants: (a) real part and (b) imaginary part.

then the independently fading components provide diversity that can be exploited by a rake receiver. This aspect of frequency-selective fading is analyzed in Section 2.5.

It is conceptually useful to define the *coherence bandwidth* as

$$B_c = \frac{1}{T_d} \tag{1-52}$$

Let B_m denote the bandwidth of $m(t)$. In general, $B_m \geq 1/T_s$ for practical modulations so flat fading occurs if $B_m << B_c$. Frequency-selective fading requires $B_m > B_c$.

To illustrate frequency-selective fading, consider the reception of a tone at frequency f_0 with two multipath components so that $m(t) = 1$ and $L_s = 2$ in (1-50). It then follows that the complex envelope has magnitude

$$|r_1(t)| = a_1 \left[1 + \left(\frac{a_2}{a_1}\right)^2 + 2\left(\frac{a_2}{a_1}\right)\cos 2\pi f_0 T_d\right]^{1/2} \tag{1-53}$$

where $T_d = \tau_1 - \tau_2$. If another tone at frequency f_1 is received, then this equation is valid with f_1 substituted for f_0. Thus, the two complex envelopes can differ considerably as f_0 and f_1 both range over a spectral band with bandwidth equal to the coherence bandwidth. If $a_1 = a_2$, then the difference between the two complex-envelope magnitudes varies from 0 to $2a_1$.

1.6 Fading Rate and Fade Duration

The *fading rate* is the rate at which the envelope of a received fading signal crosses below a specified level. Consider a time interval over which the fading parameters are constant. For a level $r \geq 0$, isotropic scattering, and Ricean fading, it can be shown that the fading rate is [1]

$$f_r = \sqrt{2\pi(\kappa+1)}\, f_d \rho \exp[-\kappa - (\kappa+1)\rho^2] I_0\left(2\rho\sqrt{\kappa(\kappa+1)}\right) \tag{1-54}$$

where κ is the Rice factor and

$$\rho = \frac{r}{\sigma_r\sqrt{2(\kappa+1)}} \tag{1-55}$$

For Rayleigh fading, $\kappa = 0$ and (1-54) becomes

$$f_r = \frac{\sqrt{\pi} f_d r}{\sigma_r} \exp(-r^2/2\sigma_r^2) \tag{1-56}$$

Equations (1-54) and (1-56) indicate that the fading rate is proportional to the Doppler spread f_d. Thus, slow fading occurs when the Doppler spread is small, whereas fast fading occurs when the Doppler spread is large.

Let T_f denote the average envelope fade duration. If the time-varying envelope is assumed to be a stationary ergodic process, then the product $f_r T_f$ is equal to $F_\alpha(r)$, the probability that the envelope is below or equal to the level r. Thus,

$$T_f = \frac{F_\alpha(r)}{f_r} \tag{1-57}$$

If the envelope has the Rice distribution, then integrating (1-28) and using (1-54), (1-55), (1-57), and (A-15), we obtain

$$T_f = \frac{1 - Q_1\left(\sqrt{2\kappa}, \sqrt{2(\kappa+1)}\rho\right)}{\sqrt{2\pi(\kappa+1)} f_d \rho \exp[-\kappa - (\kappa+1)\rho^2] I_0(2\rho\sqrt{\kappa(\kappa+1)})} \tag{1-58}$$

For Rayleigh fading, (1-56), (1-57), and the integration of (1-20) yields

$$T_f = \frac{\exp(r^2/2\sigma_r^2) - 1}{\sqrt{\pi} f_d r/\sigma_r} \tag{1-59}$$

For both Ricean and Rayleigh fading, the fade duration is inversely proportional to f_d.

2. Diversity for Fading Channels

Diversity combiners for fading channels are designed to combine independently fading copies of the same signal in different branches. The combining is done in such a way that the combiner output has a power level that varies much more slowly than that of a single copy. Although useless in improving communications over the additive-white-Gaussian-noise (AWGN) channel, diversity improves communications over fading channels because the diversity gain is large enough to overcome the noncoherent combining loss. Diversity may be provided by signal redundancy that arises in a number of different ways. *Time diversity* is provided by channel coding or by signal copies that differ in time delay. *Frequency diversity* may be available when signal copies using different carrier frequencies experience independent or weakly correlated fading. If each signal copy is extracted from the output of a separate antenna in an antenna array, then the diversity is called *spatial diversity*. *Polarization diversity* may be obtained by using two cross-polarized antennas at the same site. Although this configuration provides compactness, it is not as potentially effective as spatial diversity because the received horizontal component of an electric field is usually much weaker than the vertical component.

In the subsequent analysis, spatial diversity is emphasized, but most of the results pertain to other types of diversity as well. The three most common types of diversity combining are selective, maximal-ratio, and equal-gain combining. The last two methods use linear combining with variable weights behind each antenna. Since they usually must eventually adjust their weights, maximal-ratio and equal-gain combiners can be viewed as types of adaptive arrays. They differ from other adaptive arrays in that they are not designed to cancel interference signals.

2.1 Optimal Array

Consider an antenna array that feeds L branches. Each branch input is translated to baseband, and then either the baseband signal is applied to a matched filter and sampled or the sampled complex envelope is extracted (Appendix B-3). Alternatively, each branch input is translated to an intermediate frequency, and the sampled analytic signal is extracted. The subsequent analysis is valid for any of these types of branch processing. It is simplest to assume that the branch outputs are sampled complex envelopes. The branch outputs provide the inputs to a linear combiner. Let $\mathbf{x}(l)$ denote the discrete-time vector of the L complex-valued combiner inputs, where the index denotes the sample number. This vector can be decomposed as

$$\mathbf{x}(l) = \mathbf{s}(l) + \mathbf{n}(l) \tag{2-1}$$

where $\mathbf{s}(l)$ and $\mathbf{n}(l)$ are the discrete-time vectors of the desired signal and the interference plus thermal noise, respectively. Let $\mathbf{W}$ denote the weight vector of a linear combiner applied to the input vector. The combiner output is

$$y(l) = \mathbf{W}^T\mathbf{x}(l) = y_s + y_n \tag{2-2}$$

where T denotes the transpose of a matrix or vector,

$$y_s(l) = \mathbf{W}^T\mathbf{s}(l) \tag{2-3}$$

is the output component due to the desired signal, and

$$y_n(l) = \mathbf{W}^T\mathbf{n}(l) \tag{2-4}$$

is the output component due to the interference plus noise. The components of both $\mathbf{s}(l)$ and $\mathbf{n}(l)$ are modeled as discrete-time jointly wide-sense-stationary processes. The correlation matrix of the desired signal is defined as

$$\mathbf{R}_{ss} = E\left[\mathbf{s}^*(l)\mathbf{s}^T(l)\right] \tag{2-5}$$

and the correlation matrix of the interference plus noise is defined as

$$\mathbf{R}_{nn} = E\left[\mathbf{n}^*(l)\mathbf{n}^T(l)\right] \tag{2-6}$$

The desired-signal power at the output is

$$p_{so} = \frac{1}{2}E\left[|y_s(l)|^2\right] = \frac{1}{2}\mathbf{W}^H\mathbf{R}_{ss}\mathbf{W} \tag{2-7}$$

where the superscript H denotes the conjugate transpose. The interference plus noise power at the output is

$$p_n = \frac{1}{2}E\left[|y_n(l)|^2\right] = \frac{1}{2}\mathbf{W}^H\mathbf{R}_{nn}\mathbf{W} \tag{2-8}$$

The signal-to-interference-plus-noise ratio (SINR) at the combiner output is

$$\rho_0 = \frac{p_{so}}{p_n} = \frac{\mathbf{W}^H\mathbf{R}_{ss}\mathbf{W}}{\mathbf{W}^H\mathbf{R}_{nn}\mathbf{W}} \tag{2-9}$$

The definitions of $\mathbf{R}_{ss}$ and $\mathbf{R}_{nn}$ ensure that these matrices are Hermitian and nonnegative definite. Consequently, these matrices have complete sets of orthonormal eigenvectors, and their eigenvalues are real-valued and nonnegative. The noise power is assumed to be positive. Therefore, $\mathbf{R}_{nn}$ is positive definite and has positive eigenvalues. Since $\mathbf{R}_{nn}$ can be diagonalized, it can be expressed as [4]

$$\mathbf{R}_{nn} = \sum_{i=1}^{L} \lambda_i \mathbf{e}_i \mathbf{e}_i^H \tag{2-10}$$

where λ_i is an eigenvalue and $\mathbf{e}_i$ is the associated eigenvector.

To derive the weight vector that maximizes the SINR with no restriction on $\mathbf{R}_{ss}$, we define the Hermitian matrix

$$\mathbf{A} = \sum_{i=1}^{L} \sqrt{\lambda_i} \mathbf{e}_i \mathbf{e}_i^H \tag{2-11}$$

where the positive square root is used. Direct calculations verify that

$$\mathbf{R}_{nn} = \mathbf{A}^2 \tag{2-12}$$

and the inverse of $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \sum_{i=1}^{L} \frac{1}{\sqrt{\lambda_i}} \mathbf{e}_i \mathbf{e}_i^H \tag{2-13}$$

The matrix $\mathbf{A}$ specifies an invertible transformation of $\mathbf{W}$ into the vector

$$\mathbf{V} = \mathbf{A}\mathbf{W} \tag{2-14}$$

We define the Hermitian matrix

$$\mathbf{C} = \mathbf{A}^{-1} \mathbf{R}_{ss} \mathbf{A}^{-1} \tag{2-15}$$

Then (2-9), (2-12), (2-14), and (2-15) indicate that the SINR can be expressed as

$$\rho_0 = \frac{\mathbf{V}^H \mathbf{C} \mathbf{V}}{\|\mathbf{V}\|^2} \tag{2-16}$$

where $\| \ \|$ denotes the Euclidean norm of a vector and $\|\mathbf{V}\|^2 = \mathbf{V}^H \mathbf{V}$. Equation (2-16) is a Rayleigh quotient [4], which is maximized by $\mathbf{V} = \eta \mathbf{u}$, where $\mathbf{u}$ is the eigenvector of $\mathbf{C}$ associated with its largest eigenvalue $l_{\max}$, and η is an arbitrary constant. Thus, the maximum value of ρ_0 is

$$\rho_{\max} = l_{\max} \tag{2-17}$$

From (2-14) with $\mathbf{V} = \eta \mathbf{u}$, it follows that the *optimal weight vector that maximizes the SINR* is

$$\mathbf{W}_0 = \eta \mathbf{A}^{-1} \mathbf{u} \tag{2-18}$$

The purpose of an adaptive-array algorithm is to adjust the weight vector to converge to the optimal value, which is given by (2-18) when the maximization of the SINR is the performance criterion.

When the discrete-time dependence of $\mathbf{s}(l)$ is the same for all its components, (2-18) can be made more explicit. Let $s(l)$ denote the discrete-time sampled complex envelope of the desired signal at a fixed reference point in space. It is assumed that the desired signal is sufficiently narrowband that the difference between the maximum and minimum delays is much less than the reciprocal of the signal bandwidth. Thus, the desired-signal

copies at all the antennas are nearly aligned in time, and the desired-signal input vector may be represented as

$$\mathbf{s}(l) = s(l)\mathbf{S}_0 \tag{2-19}$$

where the *steering vector* is

$$\mathbf{S}_0 = [\alpha_1 \exp(j\,\theta_1)\ \ \alpha_2 \exp(j\,\theta_2)\ \ \ldots\ \alpha_L \exp(j\,\theta_L)]^T \tag{2-20}$$

For Rayleigh fading, the phases $\theta_i, i = 1, 2, \ldots, L$, are modeled as random variables uniformly distributed over $[0, 2\pi)$, and the attenuations $\alpha_i, i = 1, 2, \ldots, L$, have Rayleigh distribution functions, as explained in Section 1.3.

Alternatively, (2-20) can serve as a model for a narrowband desired signal that arrives at an antenna array as a plane wave and does not experience fading. Let $T_i, i = 1, 2, \ldots, L$, denote the arrival-time delay of the desired signal at the output of antenna i relative to a fixed reference point in space. Equations (2-19) and (2-20) are valid with $\theta_i = -2\pi f_c T_i, i = 1, 2, \ldots, L$, where f_c is the carrier frequency of the desired signal. The α_i, $i = 1, 2, \ldots, L$, depend on the relative antenna patterns and propagation losses. If they are all equal, then the common value can be subsumed into $s(l)$. It is convenient to define the origin of a Cartesian coordinate system to coincide with the fixed reference point. Let (x_i, y_i) denote the coordinates of antenna i. If a single plane wave arrives from direction ψ relative to the normal to the array, then

$$\theta_i = \frac{2\pi}{c} f_c(x_i \sin\psi + y_i \cos\psi), \quad i = 1, 2, \ldots, L \tag{2-21}$$

where c is the speed of an electromagnetic wave.

The substitution of (2-19) into (2-5) yields

$$\mathbf{R}_{ss} = 2p_s\mathbf{S}_0^*\mathbf{S}_0^T \tag{2-22}$$

where

$$p_s = \frac{1}{2}E[|s(l)|^2] \tag{2-23}$$

After substituting (2-22) into (2-15), it is observed that $\mathbf{C}$ may be factored:

$$\mathbf{C} = 2p_s\mathbf{A}^{-1}\mathbf{S}_0^*\mathbf{S}_0^T\mathbf{A}^{-1} = \mathbf{F}\mathbf{F}^H \tag{2-24}$$

where

$$\mathbf{F} = \sqrt{2p_s}\mathbf{A}^{-1}\mathbf{S}_0^* \tag{2-25}$$

This factorization explicitly shows that $\mathbf{C}$ is a rank-one matrix. Therefore, an eigenvector of $\mathbf{C}$ associated with the only nonzero eigenvalue is

$$\mathbf{u} = \mathbf{F} = \sqrt{2p_s}\mathbf{A}^{-1}\mathbf{S}_0^* \tag{2-26}$$

and the nonzero eigenvalue is

$$l_{\max} = \|\mathbf{F}\|^2 \tag{2-27}$$

Substituting (2-26) into (2-18), using (2-12), and then merging $\sqrt{2p_s}$ into the arbitrary constant, we obtain the *optimal weight vector for a narrowband desired signal*:

$$\mathbf{W}_0 = \eta \mathbf{R}_{nn}^{-1} \mathbf{S}_0^* \tag{2-28}$$

where η is an arbitrary constant. The maximum value of the SINR, obtained from (2-17), (2-27), (2-25), and (2-12), is

$$\rho_{\max} = 2p_s \mathbf{S}_0^T \mathbf{R}_{nn}^{-1} \mathbf{S}_0^* \ . \tag{2-29}$$

2.2 Maximal-Ratio Combining

Suppose that the interference plus noise in an antenna is zero-mean and uncorrelated with the interference plus noise in any of the other antennas in the array. Then the correlation matrix $\mathbf{R}_{nn}$ is diagonal. If the ith diagonal element has the value

$$2\sigma_i^2 = E[|n_i|^2] \tag{2-30}$$

then (2-28) implies that the weight vector that maximizes the SINR is

$$\mathbf{W}_m = \eta \left[\frac{S_{01}^*}{\sigma_1^2} \ \frac{S_{02}^*}{\sigma_2^2} \ \cdots \ \frac{S_{0N}^*}{\sigma_N^2} \right]^T \tag{2-31}$$

Equations (2-29) and (2-20) yield

$$\rho_{\max} = \sum_{i=1}^{L} \frac{p_s}{\sigma_i^2} \alpha_i^2 \tag{2-32}$$

where each term is the SINR at the output of an antenna. Linear combining that uses $\mathbf{W}_m$ is called *maximal-ratio combining* (MRC). It is optimal only if the interference-plus-noise signals in all the diversity branches are uncorrelated. As discussed subsequently, the maximal-ratio combiner can also be derived as the maximum-likelihood estimator associated with a multivariate Gaussian density function. The critical assumption in the derivation is that the noise process in each array branch is both Gaussian and independent of the noise processes in the other branches.

In most applications, the interference-plus-noise power in each array branch is approximately equal, and it is assumed that $\sigma_i^2 = \sigma^2$, $i = 1, 2, \ldots, L$. If this common value is merged with the constant in (2-28) or (2-31), then the MRC weight vector is

$$\mathbf{W}_m = \eta \mathbf{S}_0^* \tag{2-33}$$

and the corresponding maximum SINR is

$$\rho_{\max} = \frac{p_s}{\sigma^2} \sum_{i=1}^{L} \alpha_i^2 \tag{2-34}$$

Since the weight vector is not a function of the interference parameters, the combiner attempts no interference cancellation. The interference signals are ignored while the combiner does coherent combining of the desired signal. Equations (2-3), (2-33), (2-19), and (2-20) yield the desired part of the combiner output:

$$y_s(l) = \mathbf{W}_m^T \mathbf{s}(l) = \eta s(l) \sum_{i=1}^{L} \alpha_i^2 \tag{2-35}$$

Since $y_s(l)$ is proportional to $s(l)$, the MRC equalizes the phases of the signal copies in the array branches, a process called *cophasing*. If cophasing can be done rapidly enough to be practical, then so can coherent demodulation.

If each $\alpha_i, i = 1, 2, \ldots, L$, is modeled as a random variable with an identical probability distribution function, then (2-34) implies that

$$E[\rho_{\max}] = L \frac{p_s}{\sigma^2} E[\alpha_1^2] \tag{2-36}$$

which indicates a gain in the mean SINR that is proportional to L.

There are several ways to implement cophasing [5]. Unlike most other cophasing systems, the *phase stripper* does not require a pilot signal. Figure 7 depicts branch k of a digital version of a maximal-ratio combiner with a phase stripper. It is assumed that the interference-plus-noise power in each branch is equal so that only cophasing and amplitude multiplication are required for the MRC. In the absence of noise, the angle-modulated input signal is assumed to have the form

$$c_k(l) = \alpha_k s(l) \exp[j\theta_k] = \alpha_k \exp[j\phi(l) + j\theta_k] \tag{2-37}$$

where α_k is the amplitude, $\phi(l)$ is the angle modulation carried by all the signal copies in the diversity branches, and θ_k is the undesired phase shift in branch k, which is assumed to be constant for at least two consecutive

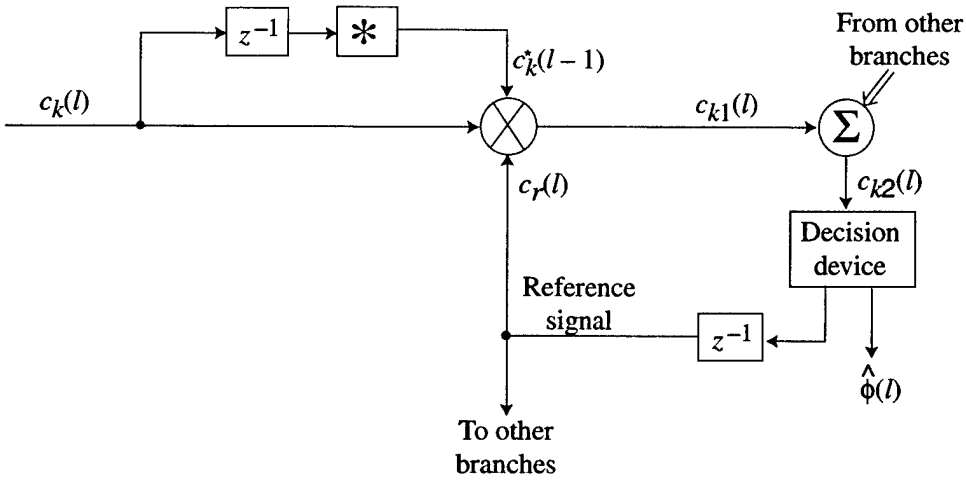


Figure 7. Branch k of a maximal-ratio combiner with a phase stripper.

samples. The signal $c_k^*(l-1)$ is produced by a delay and complex conjugation. During steady-state operation following an initialization process, the reference signal is assumed to have the form

$$c_r(l) = \exp[j\phi(l-1) + j\psi] \tag{2-38}$$

where ψ is a phase angle. The three signals $c_r(l)$, $c_k(l)$, and $c_k^*(l-1)$ are multiplied together to produce

$$c_{k1}(l) = \alpha_k^2 \exp[j\phi(l) + j\psi] \tag{2-39}$$

which as been stripped of the undesired phase shift θ_k. This signal is combined with similar signals from the other diversity branches that use the same reference signal. The input to the decision device is

$$c_{k2}(l) = \sum_{k=1}^{L} \alpha_k^2 \exp[j\phi(l) + j\psi] = e^{j\psi} s(l) \sum_{k=1}^{L} \alpha_k^2 \tag{2-40}$$

which indicates that MRC has been obtained by phase equalization, as in (2-35). After extracting the phase $\phi(l) + \psi$, the decision device produces the demodulated sequence $\hat{\phi}(l)$, which is an estimate of $\phi(l)$, by some type of phase-recovery loop [6]. The device also produces the complex exponential $\exp[j\phi(l) + j\psi]$. After a delay, the complex exponential provides the reference signal of (2-38).

2.3 Bit Error Probabilities for Coherent Binary Modulations

Suppose that the desired-signal modulation is binary phase-shift keying (PSK). A transmitted bit is equally likely to be a 0 or a 1 and is represented by $+\psi(t)$ or $-\psi(t)$, respectively. Each received signal copy in a diversity branch experiences independent Rayleigh fading that is constant during the signal interval. The received signal in branch i is

$$r_i(t) = \text{Re}[\alpha_i e^{j\theta_i} x\psi(t) e^{j2\pi f_c t}] + n_i(t), \quad 0 \le t \le T, \quad i = 1, 2, \ldots, L \tag{2-41}$$

where $x = +1$ or -1 depending on the transmitted bit, each α_i is an amplitude, each θ_i is a phase shift, f_c is the carrier frequency, T is the signal duration, and $n_i(t)$ is the noise. It is assumed that either the interference is absent or, more generally, that the received interference plus noise in each diversity branch can be modeled as independent, zero-mean, white Gaussian noise with the same two-sided power spectral density $N_0/2$.

Although MRC maximizes the SINR after linear combining, the theory of maximum-likelihood detection (cf. Section 2.6) is needed to determine an optimal decision variable that can be compared to a threshold. The initial branch processing before sampling could entail extraction of the complex envelope (Appendix B-3), passband matched-filtering followed by a frequency translation to baseband, or, equivalently, a translation followed by baseband matched-filtering [6]. Since it is slightly simpler, we assume the

latter in this analysis. The same results are obtained if one assumes the extraction of the complex envelope and uses the equations of Appendix B-3.

Using $2\text{Re}(x) = x + x^*$ and discarding a negligible integral, it is found that after the translation to baseband, the matched filter in each diversity branch, which is matched to $\psi(t)$, produces the samples

$$\begin{aligned} y_i &= \int_0^T 2r_i(t)e^{-j2\pi f_c t}\psi^*(t)dt \\ &= 2\mathcal{E}\alpha_i e^{j\theta_i}x + \int_0^T 2n_i(t)e^{-j2\pi f_c t}\psi^*(t)dt, \quad i = 1, 2, \ldots, L \end{aligned} \tag{2-42}$$

where a factor of "2" has been inserted for analytical convenience, and the desired-signal energy per bit in the absence of fading and multipath is

$$\mathcal{E} = \frac{1}{2}\int_0^T |\psi(t)|^2 dt \tag{2-43}$$

These samples provide sufficient statistics that contain all the relevant information in the received signal copies in the L diversity branches.

It is assumed that $\psi(t)$ has a spectrum confined to $|f| < f_c$. The white Gaussian noise process $n_i(t)$ has autocorrelation

$$E[n_i(t)n_i(t+\tau)] = \frac{N_0}{2}\delta(\tau) \tag{2-44}$$

where $\delta(\tau)$ is the Dirac delta function. Using $2\text{Re}(x) = x + x^*$, $2\text{Im}(x) = x - x^*$, the spectral limitations of $\psi(t)$, (2-43), and (2-44), we find that the noise term in (2-42) has independent real and imaginary components with the same variance $2\mathcal{E}N_0$. Therefore, given x, α_i, and θ_i, the branch likelihood function or conditional probability density function of y_i is

$$f(y_i|x, \alpha_i, \theta_i) = \frac{1}{4\pi\mathcal{E}N_0}\exp\left[-\frac{|y_i - 2\mathcal{E}\alpha_i e^{j\theta_i}x|^2}{4\mathcal{E}N_0}\right], \quad i = 1, 2, \ldots, L \tag{2-45}$$

Since the branch samples are statistically independent, the log-likelihood function for the vector $\mathbf{y} = (y_1\, y_2 \ldots y_L)$ given the $\{\alpha_i\}$ and the $\{\theta_i\}$ is

$$\ln[f(\mathbf{y}|x, \boldsymbol{\alpha}, \boldsymbol{\theta})] = \sum_{i=1}^{L} \ln[f(y_i|x, \alpha_i, \theta_i)] \tag{2-46}$$

The receiver decides in favor of a 0 or 1 depending on whether $x = +1$ or $x = -1$ gives the larger value of the log-likelihood function. Substituting (2-45) into (2-46), eliminating irrelevant terms and factors that do not depend on the value of x, we find that the maximum-likelihood decoder can base its decision on the single variable

$$\begin{aligned} U &= \sum_{i=1}^{L} \text{Re}\left(\alpha_i e^{-j\theta_i} y_i\right) \\ &= 2\mathcal{E}\sum_{i=1}^{L}\alpha_i^2 + \sum_{i=1}^{L}\alpha_i N_i \end{aligned} \tag{2-47}$$

where, given the value of θ_i, N_i is the real-valued, zero-mean, Gaussian random variable

$$N_i = \text{Re}\left[e^{-j\theta_i}\int_0^T 2n_i(t)e^{-j2\pi f_c t}\psi^*(t)dt\right] \tag{2-48}$$

If we let $\mathbf{s}(l) = [y_1\ y_2\ \cdots\ y_L]^T$ and use (2-33), we obtain $U = \text{Re}[\mathbf{W}_m^T\mathbf{s}(l)]$. Since taking the real part of $\mathbf{W}_m^T\mathbf{s}(l)$ serves only to eliminate orthogonal noise, the decision variable U is produced by baseband MRC. The decision variable is compared to a threshold equal to zero to determine the bit state.

Since (2-47) is computed in either case, the implementation of the maximum-likelihood decoder may use either maximal-ratio *predetection combining* before the demodulation, as illustrated in Figure 8(a), or *postdetection combining* following the demodulation, as illustrated in Figure 8(b). Since the optimal coherent matched-filter or correlation demodulator performs a linear operation on the $\{y_i\}$, both predetection and postdetection combining provide the same decision variable, and hence the same performance.

If the $\{\alpha_i\}$ and $\{\theta_i\}$ are given, then the decision variable has a Gaussian distribution with mean

$$E(U) = 2\mathcal{E}\sum_{i=1}^{L}\alpha_i^2 \tag{2-49}$$

Since the $\{n_i(t)\}$ and, hence, the $\{N_i\}$ are independent, the variance of U is

$$\sigma_u^2 = \sum_{i=1}^{L}\alpha_i^2\ \text{var}(N_i) \tag{2-50}$$

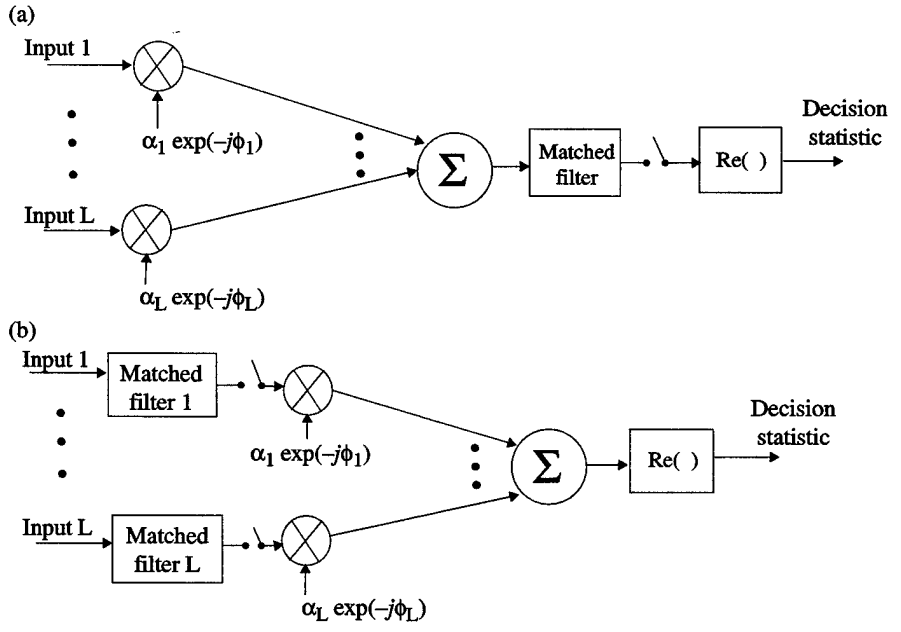


Figure 8. Maximal-ratio combiner for PSK with (a) predetection combining and (b) postdetection combining. Coherent equal-gain combiner for PSK omits the factors $\{\alpha_i\}$.

The variance of N_i can be evaluated from (2-43), (2-44), and (2-48). It then follows from (2-50) that

$$\sigma_u^2 = 2\mathcal{E}N_0 \sum_{i=1}^{L} \alpha_i^2 \tag{2-51}$$

Because of the symmetry, the bit error probability can be determined by assuming that $x = +1$, corresponding to a transmitted 0. A decision error is made if $U < 0$. Since the decision variable has a Gaussian conditional distribution and neither $E(U)$ nor σ_u^2 depends on the $\{\theta_i\}$, a standard derivation indicates that the conditional bit error probability given the $\{\alpha_i\}$ is

$$P_{b|\alpha}(\gamma_b) = Q(\sqrt{2\gamma_b}) \tag{2-52}$$

where

$$\gamma_b = \frac{\mathcal{E}}{N_0} \sum_{i=1}^{L} \alpha_i^2 \tag{2-53}$$

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy = \frac{1}{2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) \tag{2-54}$$

and erfc( ) is the complementary error function. The bit error probability is determined by averaging $P_{b|\alpha}(\gamma_b)$ over the distribution of γ_b, which depends on the $\{\alpha_i\}$ and embodies the statistics of the fading channel.

Suppose that independent Rayleigh fading occurs so that each of the $\{\alpha_i\}$ is independent with the identical Rayleigh distribution and $E[\alpha_i^2] = E[\alpha_1^2]$. As shown in Appendix A-4, α_i^2 is exponentially distributed. Therefore, γ_b in (2-53) is the sum of L independent, identically and exponentially distributed random variables. From (A-49), it follows that the probability density function of γ_b is

$$f_\gamma(x) = \frac{1}{(L-1)!\bar{\gamma}^L} x^{L-1} \exp\left(-\frac{x}{\bar{\gamma}}\right) u(x) \tag{2-55}$$

where the average signal-to-noise ratio (SNR) per branch is

$$\bar{\gamma} = \frac{\mathcal{E}}{N_0} E[\alpha_1^2] \tag{2-56}$$

The bit error probability is determined by averaging (2-52) over the density given by (2-55). Thus,

$$P_b(L) = \int_0^\infty Q(\sqrt{2x}) \frac{1}{(L-1)!\bar{\gamma}^L} x^{L-1} \exp\left(-\frac{x}{\bar{\gamma}}\right) dx \tag{2-57}$$

Direct calculations verify that since L is an integer,

$$\frac{d}{dx} Q\left(\sqrt{2x}\right) = -\frac{1}{2\sqrt{\pi}} \frac{\exp(-x)}{\sqrt{x}} \tag{2-58}$$

$$\frac{d}{dx}\left[e^{-x/\bar{\gamma}} \sum_{i=0}^{L-1} \frac{(x/\bar{\gamma})^i}{i!}\right] = -\frac{1}{(L-1)!\bar{\gamma}^L} x^{L-1} \exp\left(-\frac{x}{\bar{\gamma}}\right) \tag{2-59}$$

Applying integration by parts to (2-57), using (2-58), (2-59), and $Q(0) = 1/2$, we obtain

$$P_b(L) = \frac{1}{2} - \sum_{i=0}^{L-1} \frac{1}{i!\bar{\gamma}^i 2\sqrt{\pi}} \int_0^\infty \exp\left[-x\left(1+\bar{\gamma}^{-1}\right)\right] x^{i-1/2}\, dx \qquad (2\text{-}60)$$

This integral can be evaluated in terms of the gamma function, which is defined in (A-12). A change of variable in (2-60) yields

$$P_b(L) = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{\bar{\gamma}}{1+\bar{\gamma}}} \sum_{i=0}^{L-1} \frac{\Gamma(i+1/2)}{\sqrt{\pi} i!(1+\bar{\gamma})^i} \qquad (2\text{-}61)$$

Since $\Gamma(1/2) = \sqrt{\pi}$, the bit error probability for no diversity or a single branch is

$$p = P_b(1) = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}}{1+\bar{\gamma}}}\right) \qquad \text{(PSK, QPSK)} \qquad (2\text{-}62)$$

Since $\Gamma(x) = (x-1)\Gamma(x-1)$, it follows that

$$\Gamma(k+1/2) = \frac{\sqrt{\pi}\,\Gamma(2k)}{2^{2k-1}\Gamma(k)} = \frac{\sqrt{\pi} k!}{2^{2k-1}} \binom{2k-1}{k}, \quad k \geq 1 \qquad (2\text{-}63)$$

Solving (2-62) to determine $\bar{\gamma}$ as a function of p and then using this result and (2-63) in (2-61) gives

$$P_b(L) = p - (1-2p) \sum_{i=1}^{L-1} \binom{2i-1}{i} [p(1-p)]^i \qquad (2\text{-}64)$$

This expression explicitly shows the change in the bit error probability as the number of diversity branches increases. Equations (2-62) and (2-64) are valid for quadriphase-shift keying (QPSK) because the latter can be transmitted as two independent binary PSK waveforms in phase quadrature.

An alternative expression for $P_b(L)$, which may be obtained by a far more complicated calculation entailing the use of the properties of the Gauss hypergeometric function, is [3], [7]

$$P_b(L) = p^L \sum_{i=0}^{L-1} \binom{L+i-1}{i} (1-p)^i \qquad (2\text{-}65)$$

By using mathematical induction, this equation can be derived from (2-64) without invoking the hypergeometric function.

From a known identity for the sum of binomial coefficients [8], it follows that

$$\sum_{i=0}^{L-1} \binom{L+i-1}{i} = \binom{2L-1}{L} \qquad (2\text{-}66)$$

Since $1-p \leq 1$, (2-65) and (2-66) imply that

$$P_b(L) \leq \binom{2L-1}{L} p^L \tag{2-67}$$

This upper bound becomes tighter as $p \to 0$. If $\bar{\gamma} >> 1$ so that $p << 1$, (2-62) implies that $p \approx 1/4\bar{\gamma}$ and (2-67) indicates that the bit error probability decreases inversely with $\bar{\gamma}^L$, thereby demonstrating the large performance improvement provided by diversity.

The advantage of MRC is critically dependent on the assumption of uncorrelated fading in each diversity branch. If there is complete correlation so that the $\{\alpha_i\}$ are all equal and the fading occurs simultaneously in all the diversity branches, then $\gamma_b = L\mathcal{E}\alpha_1^2/N_0$. Therefore, γ_b has a chi-square distribution with 2 degrees of freedom and probability density function

$$f_\gamma^c(x) = \frac{1}{L\bar{\gamma}} \exp\left(-\frac{x}{L\bar{\gamma}}\right) u(x) \tag{2-68}$$

where $\bar{\gamma}$ is defined by (2-56) and the superscript c denotes correlated fading. A derivation similar to that of (2-61) yields

$$P_b^c(L) = \frac{1}{2}\left(1 - \sqrt{\frac{L\bar{\gamma}}{1+L\bar{\gamma}}}\right) \quad \text{(PSK, QPSK)} \tag{2-69}$$

When $L\bar{\gamma} >> 1$,

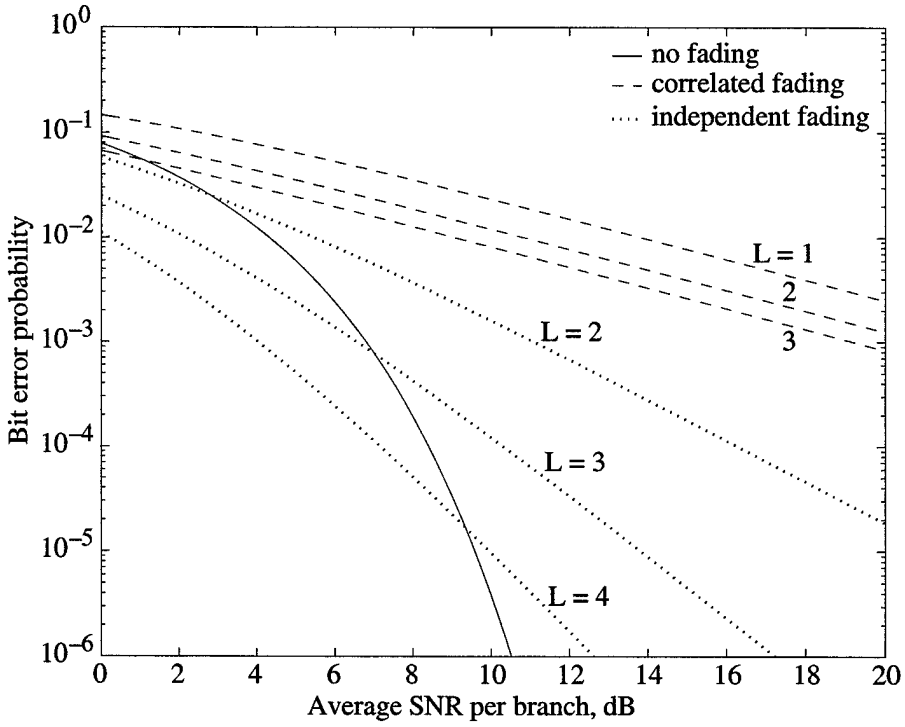
$$P_b^c(L) \approx \frac{1}{4L\bar{\gamma}} \approx \frac{p}{L} \quad , \quad p << 1 \quad \text{(PSK, QPSK)} \tag{2-70}$$

where p is given by (2-62). A comparison of (2-70) with (2-67) shows the large disparity in performance between a system with completely correlated fading and one with uncorrelated fading.

Plots of the bit error probability for a single branch with no fading, L branches with independent fading and MRC, and L branches with completely correlated fading and MRC are shown in Figure 9. Equations (2-52), (2-62), (2-64), and (2-69) are used in generating the plots. The independent variable is the average SNR per branch for a bit, which is equal to $\bar{\gamma}$ for MRC and is equal to $\gamma_b = \mathcal{E}/N_0$ for the single branch with no fading. The average SNR per bit for MRC is $L\bar{\gamma}$. The figure demonstrates the advantage of diversity combining and independent fading.

For multiple frequency-shift keying (MFSK), one of q equal-energy orthogonal signals $s_1(t), s_2(t), \ldots, s_q(t)$, each representing $\log_2 q$ bits, is transmitted. The maximum-likelihood decoder generates q decision variables corresponding to the q possible nonbinary symbols. The decoder decides in favor of the symbol associated with the largest of the decision variables. Matched filters for the q orthogonal signals are needed in every diversity branch. Because of the orthogonality, each filter matched to $s_k(t)$ has a zero response to $s_l(t)$, $l \neq k$, at the sampling time. When symbol l represented

Figure 9. Bit error probability of PSK for no fading, completely correlated fading, and independent fading.



by $s_l(t)$ is received in the presence of white Gaussian noise, matched-filter k of branch i produces the sample

$$y_{ki} = 2\mathcal{E}\alpha_i e^{j\theta_i}\delta_{kl} + \int_0^T 2n_i(t)e^{-j2\pi f_c t}s_k^*(t)\,dt\,, \tag{2-71}$$

where $\delta_{kl} = 1$ if $k = l$ and $\delta_{kl} = 0$ if $k \neq l$ and

$$\mathcal{E} = \frac{1}{2}\int_0^T |s_k(t)|^2\,dt\,, \quad k = 1, 2, \ldots, q \tag{2-72}$$

It is assumed that each $s_k(t)$ has a spectrum confined to $|f| < f_c$. Using these spectral limitations and (2-44), we find that the real and imaginary components of the noise term in (2-71) are independent and have the same variance $2\mathcal{E}N_0$ and that the conditional probability density function of y_{ki} given the values of l, α_i, and θ_i is

$$f(y_{ki}|l, \alpha_i, \theta_i) = \frac{1}{4\pi\mathcal{E}N_0}\exp\left[-\frac{|y_{ki} - 2\mathcal{E}\alpha_i e^{j\theta_i}\delta_{kl}|^2}{4\mathcal{E}N_0}\right] \tag{2-73}$$

For coherent MFSK, the $\{\alpha_i\}$ and the $\{\theta_i\}$ are assumed to be known. Since the noise in each branch is assumed to be independent, the likelihood function is the product of qL densities given by (2-73) for $k = 1, 2, \ldots, q$ and $i = 1, 2, \ldots, L$. Forming the log-likelihood function, observing that $\sum_k \delta_{kl}^2 = 1$, and eliminating irrelevant terms and factors that are independent of l, we find that the maximization of the log-likelihood function is equivalent to selecting the largest of q decision variables, one for each of $s_1(t)$, $s_2(t)$, $\ldots$, $s_q(t)$. They are

$$U_l = \sum_{i=1}^{L} \mathrm{Re}\left(\alpha_i e^{-j\theta_i} y_{li}\right) \quad , \quad l = 1, 2, \ldots, q \tag{2-74}$$

Consider coherent binary frequency-shift keying (FSK). Because of the symmetry of the model, $P_b(L)$ can be calculated by assuming that $s_1(t)$ was transmitted. With this assumption, the two decision variables become

$$U_1 = 2\mathcal{E} \sum_{i=1}^{L} \alpha_i^2 + \sum_{i=1}^{L} \alpha_i N_{1i} \tag{2-75}$$

$$U_2 = \sum_{i=1}^{L} \alpha_i N_{2i} \tag{2-76}$$

where N_{1i} and N_{2i} are independent, real-valued, Gaussian noise variables given by

$$N_{ki} = \mathrm{Re}\left[e^{-j\theta_i} \int_0^T 2n_i(t) e^{-j2\pi f_c t} s_{ki}^*(t)\, dt\right] , \quad k = 1, 2 \tag{2-77}$$

A derivation similar to the one for coherent PSK indicates that (2-64) and (2-65) are again valid for coherent FSK provided that

$$p = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}}{2 + \bar{\gamma}}}\right) \qquad \text{(coherent FSK)} \tag{2-78}$$

which can also be obtained by observing the presence of two independent noise variables and, hence, substituting $\bar{\gamma}/2$ in place of $\bar{\gamma}$ in (2-62). Thus, in a fading environment, PSK retains its usual 3 dB advantage over coherent FSK.

The preceding analysis for independent Rayleigh fading can be extended to independent Nakagami fading if the parameter m is a positive integer. From (1-29) and elementary probability, it follows that the probability density function of each random variable $\gamma_i = \mathcal{E}\alpha_i^2/N_0$ is

$$f_{\gamma i}(x) = \frac{m^m}{(m-1)!\,\bar{\gamma}^m}\, x^{m-1} \exp\left(-\frac{mx}{\bar{\gamma}}\right) u(x) \, , \quad m = 1, 2, \ldots \tag{2-79}$$

where $\bar{\gamma}$ is defined by (2-56). As indicated in Appendix A-2, the characteristic function of γ_i is

$$C_{\gamma i}(j\nu) = \frac{1}{(1 - j\frac{\bar{\gamma}}{m}\nu)^m} \tag{2-80}$$

If γ_b in (2-53) is the sum of L independent, identically-distributed random variables, then it has the characteristic function

$$C_{\gamma}(j\nu) = \frac{1}{(1 - j\frac{\bar{\gamma}}{m}\nu)^{mL}} \tag{2-81}$$

The inverse of this function yields the probability density function

$$f_\gamma(x) = \frac{1}{(mL-1)!\,(\bar{\gamma}/m)^{mL}}\, x^{mL-1} \exp\left(-\frac{mx}{\bar{\gamma}}\right) u(x)\ ,\ m = 1, 2, \ldots \tag{2-82}$$

The form of this expression is the same as that in (2-55) except that L and $\bar{\gamma}$ are replaced by mL and $\bar{\gamma}/m$, respectively. Consequently, the derivation following (2-55) is valid once the replacements are made, and

$$P_b(L) = p - (1-2p) \sum_{i=1}^{mL-1} \binom{2i-1}{i} [p(1-p)]^i \tag{2-83}$$

where

$$p = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}}{m+\bar{\gamma}}}\right) \quad \text{(PSK, QPSK)} \tag{2-84}$$

$$p = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}}{2m+\bar{\gamma}}}\right) \quad \text{(coherent FSK)} \tag{2-85}$$

These results can be approximately related to Ricean fading by using (1-30). Figure 10 displays the bit error probability for Nakagami fading with $m = 4$, PSK, and $L = 1, 2, 3$, and 4 diversity branches.

2.4 Equal-Gain Combining

Coherent equal-gain combining (EGC) performs cophasing, but does not correct for unequal values of α_i/σ_i^2, $i = 1, 2, \ldots, L$, where $\alpha_i = |S_{0i}|$. Thus,

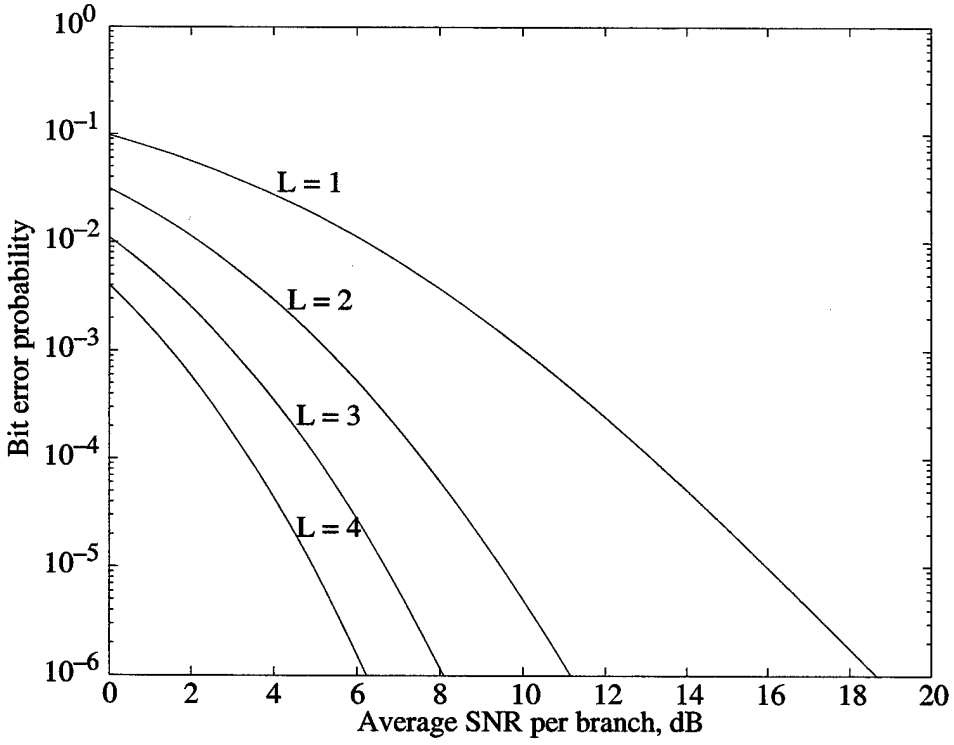


Figure 10. Bit error probability of PSK for Nakagami fading with $m = 4$.

when a narrowband desired signal experiences fading, instead of (2-31) and (2-20), the EGC weight vector is

$$\mathbf{W}_e = \eta[\exp(-j\theta_1) \quad \exp(-j\theta_2) \quad \ldots \quad \exp(-j\theta_L)]^T \tag{2-86}$$

where θ_i is the phase shift of the desired signal in branch i. When MRC is optimal and the values of the $\{\alpha_i/\sigma_i^2\}$ are unequal, EGC is suboptimal, but requires much less information about the channel. If the interference plus noise in each array branch is zero-mean and uncorrelated with the other branches and $E[|n_i|^2] = 2\sigma^2, i = 1, 2, \ldots, L$, then $\mathbf{R}_{nn}$ is diagonal, and (2-9), (2-20), and (2-22) with $\mathbf{W} = \mathbf{W}_e$ give the output SINR

$$\rho_0 = \frac{p_s}{L\sigma^2}\left(\sum_{i=1}^{L}\alpha_i\right)^2 \tag{2-87}$$

It can be verified by applying the Schwarz inequality for inner products that this SINR is less than or equal to $\rho_{\max}$ given by (2-34). Figure 8 pertains to EGC with predetection and postdetection combining if the factors $\{\alpha_i\}$ are omitted.

In a Rayleigh-fading environment, each α_i, $i = 1, 2, \ldots, L$, has a Rayleigh probability distribution function. If the desired signal in each array branch is uncorrelated with the other branches and has identical average power, then using (A-36), we obtain

$$E[\alpha_i^2] = E[\alpha_1^2], \quad E[\alpha_i] = \left\{\frac{\pi}{4}E[\alpha_1^2]\right\}^{1/2}, \quad i = 1, 2, \ldots, L \tag{2-88}$$

$$E[\alpha_i\alpha_k] = E[\alpha_i]\,E[\alpha_k] = \frac{\pi}{4}E[\alpha_1^2], \quad i \neq k \tag{2-89}$$

These equations and (2-87) give

$$E[\rho_0] = \left[1 + (L-1)\frac{\pi}{4}\right]\frac{p_s}{\sigma^2}E[\alpha_1^2] \tag{2-90}$$

which exceeds $\pi/4$ times $E[\rho_{\max}]$ given by (2-36) for MRC.

As a specific example in which MRC is suboptimal, assume that both the desired and interference signals are narrowband, do not experience fading, and arrive as plane waves. The array antennas are sufficiently close that the steering vector $\mathbf{S}_0$ of the desired signal and the steering vector $\mathbf{J}_0$ of the interference signal can be represented by

$$\mathbf{S}_0 = \left[e^{-j2\pi f_0\tau_1} \quad e^{-j2\pi f_0\tau_2} \quad \ldots \quad e^{-j2\pi f_0\tau_L}\right]^T \tag{2-91}$$

$$\mathbf{J}_0 = \left[e^{-j2\pi f_0\delta_1} \quad e^{-j2\pi f_0\delta_2} \quad \ldots \quad e^{-j2\pi f_0\delta_L}\right]^T \tag{2-92}$$

The correlation matrix for the interference plus noise is

$$\mathbf{R}_{nn} = 2p_n\,\mathbf{I} + 2p_i\,\mathbf{J}_0^*\,\mathbf{J}_0^T \tag{2-93}$$

where p_n and p_i are the noise and interference powers, respectively, in each array branch. This equation shows explicitly that the interference in one branch is correlated with the interference in the other branches. A direct matrix multiplication using $\|\mathbf{J}_0\|^2 = L$ verifies that

$$\mathbf{R}_{nn}^{-1} = \frac{1}{2p_n}\left(\mathbf{I} - \frac{g\,\mathbf{J}_0^*\mathbf{J}_0^T}{Lg+1}\right) \tag{2-94}$$

where $g = p_i/p_n$ is the interference-to-noise ratio in each array branch. After merging $1/2p_n$ with the constant in (2-28), it is found that the optimal weight vector is

$$\mathbf{W}_0 = \eta\left(\mathbf{S}_0^* - \frac{\xi L g}{Lg+1}\mathbf{J}_0^*\right) \tag{2-95}$$

where ξ is the normalized inner product

$$\xi = \frac{1}{L}\mathbf{J}_0^T\mathbf{S}_0^* \tag{2-96}$$

The corresponding maximum SINR, which is calculated by substituting (2-91), (2-94), and (2-96) into (2-29), is

$$\rho_{\max} = L\,\gamma_s\left(1 - \frac{|\xi|^2 L g}{Lg+1}\right) \tag{2-97}$$

where $\gamma_s = p_s/p_n$ is the signal-to-noise ratio in each branch. Equations (2-91), (2-92), and (2-96) indicate that $0 \le |\xi| \le 1$ and $|\xi| = 1$ if $L = 1$. Equation (2-97) indicates that $\rho_{\max}$ decreases as $|\xi|$ increases if $L \ge 2$ and is nearly directly proportional to L if $g >> 1$.

In this example, since the values of the $\{\alpha_i/\sigma_i^2\}$ are all equal, both MRC and EGC use the weight vector of (2-86) with $\theta_i = -2\pi f_c \tau_i, i = 1, 2, \ldots, L$, which gives $\mathbf{W} = \eta\mathbf{S}_0^*$. Substituting (2-22), (2-91)–(2-93), and (2-96) into (2-9) gives the SINR for MRC and EGC:

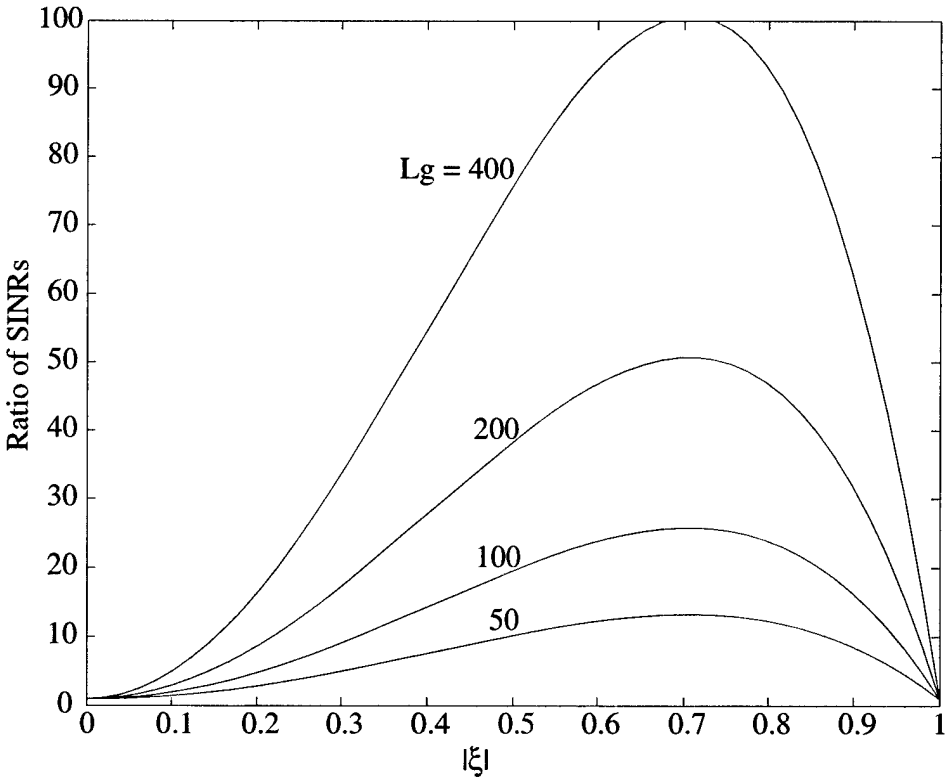
$$\rho_0 = \frac{L\,\gamma_s}{1 + |\xi|^2 L g} \tag{2-98}$$

Both $\rho_{\max}$ and ρ_0 equal $L\,\gamma_s$, the peak value, when $\xi = 0$. They both equal $L\,\gamma_s/(Lg+1)$ when $|\xi| = 1$, which occurs when both the desired and interference signals arrive from the same direction or $L = 1$. Using calculus, it is found that the maximum value of $\rho_{\max}/\rho_0$, which occurs when $|\xi| = 1/\sqrt{2}$, is

$$\left(\frac{\rho_{\max}}{\rho_0}\right)_{\max} = \frac{(Lg/2+1)^2}{Lg+1}\ , \quad L \ge 2 \tag{2-99}$$

This ratio approaches $Lg/4$ for large values of Lg. Thus, an adaptive array based on the maximization of the SINR has the potential to significantly outperform MRC or EGC if $Lg >> 1$ under the conditions of this example. Figure 11 displays $\rho_{\max}/\rho_0$ as a function of $|\xi|$ for various values of Lg.

Figure 11. Ratio of the maximum SINR to the maximal-ratio-combiner SINR.



When accurate phase estimation is unavailable so that neither cophasing nor coherent demodulation is possible, then postdetection combining following noncoherent demodulation can provide a significant performance improvement over a system with no diversity. For FSK or minimum-shift keying, postdetection combining with a frequency discriminator is illustrated in Figure 12. Each intermediate frequency (IF) is sampled, converted to a discrete-time complex baseband signal, and then demodulated by a digital frequency discriminator [9]. The square of the magnitude or possibly the magnitude of the discrete-time complex baseband signal is used to weight the output of each branch. If the noise power in each branch is approximately the same and much smaller then the desired-signal power, then this weighting is a good approximation of the weighting used in MRC, but it is suboptimal since cophasing is absent.

An alternative is postdetection EGC. However, when the desired-signal power is very low in a branch, then that branch contributes only noise to the EGC output. This problem is eliminated if each branch has a threshold device that blocks the output of that branch if the desired-signal power falls below the threshold.

A block diagram of a differential phase-shift keying (DPSK) receiver with postdetection EGC is depicted in Figure 13. For equally likely binary symbols, the error probability is the same regardless of whether two consecutive symbols are the same or different. Assuming that they are the same and that the fading is constant over two symbols, the EGC decision statistic is [3]

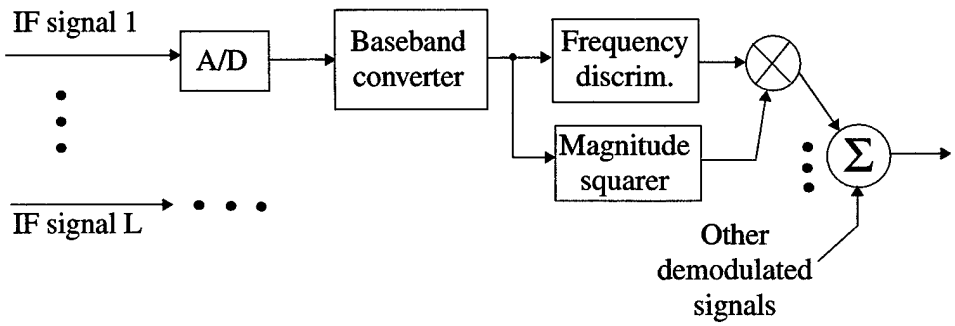


Figure 12. Postdetection combining with frequency discriminator.

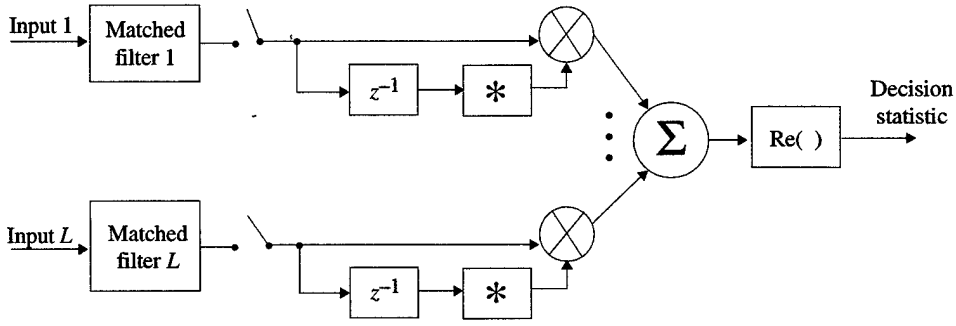


Figure 13. Equal-gain combiner for DPSK with postdetection combining.

$$U = \mathrm{Re}\left[\sum_{i=1}^{L}\left(2\mathcal{E}\alpha_i e^{j\theta_i} + N_{1i}\right)\left(2\mathcal{E}\alpha_i e^{-j\theta_i} + N_{2i}^*\right)\right] \tag{2-100}$$

where N_{1i} and N_{2i} are independent, complex-valued, Gaussian noise variables arising from two consecutive symbol intervals. A derivation [3] indicates that if the $\{\alpha_i\}$ are independent but have identical Rayleigh distributions, then $P_b(L)$ is given by (2-64), (2-65), and (2-67) with the single-branch bit error probability

$$p = \frac{1}{2(1+\bar{\gamma})} \qquad \text{(DPSK)} \tag{2-101}$$

where $\bar{\gamma}$ is given by (2-56). Equation (2-101) can be directly derived by observing that the conditional bit error probability for DPSK with no diversity is $\frac{1}{2}\exp(-\gamma_b)$ and then integrating the equation over the density (2-55) with $L = 1$. A comparison of (2-101) with (2-78) indicates that DPSK with EGC and coherent FSK with MRC give nearly the same performance in a Rayleigh-fading environment if $\bar{\gamma} >> 1$.

Figure 14 is a block diagram of a classical noncoherent MFSK receiver with postdetection square-law EGC. Each branch contains filters matched to the equal-energy orthogonal signals $s_1(t)$, $s_2(t)$, ..., $s_q(t)$. To derive this receiver from the maximum-likelihood criterion, we assume that the $\{\alpha_i\}$ and the $\{\theta_i\}$ in (2-71) are random variables. We expand the argument of the exponential function in (2-73), assume that θ_i is uniformly distributed over $[0, 2\pi)$, and integrate over the density of θ_i. The integral may be evaluated by expressing y_{ki} in polar form, using (A-30), and observing that the

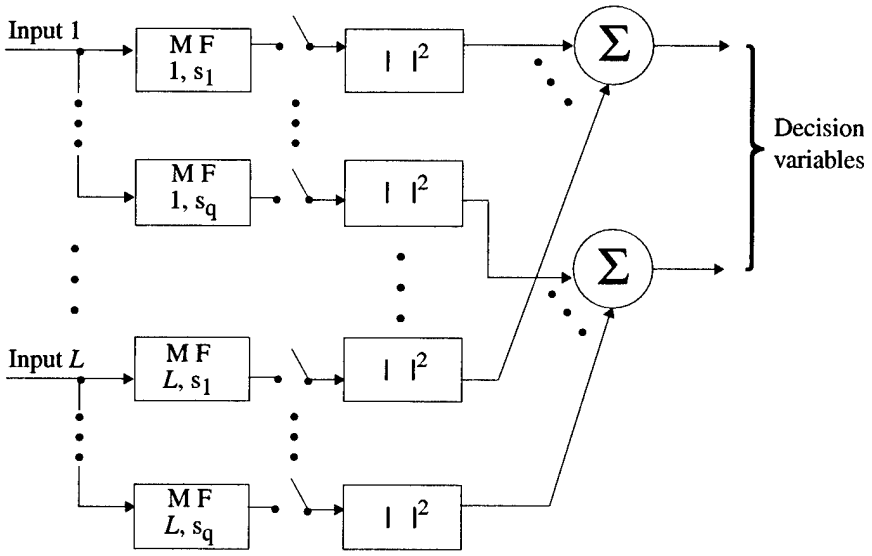


Figure 14. Equal-gain combiner for noncoherent MFSK with postdetection combining.

integral is over one period of a periodic integrand. Thus, we obtain the conditional density function

$$f(y_{ki}|l,\alpha_i) = \frac{1}{4\pi\mathcal{E}N_0}\exp\left[-\frac{|y_{ki}|^2+4\mathcal{E}\alpha_i^2\delta_{kl}}{4\mathcal{E}N_0}\right] I_0\left(\frac{\alpha_i|y_{ki}|\delta_{kl}}{N_0}\right) \qquad (2\text{-}102)$$

Assuming that α_i has the Rayleigh probability density function given by (1-20), the density $f(y_{ki}|l)$ may be evaluated by using the identity (A-33). The likelihood function is the product of qL densities for $k = 1, 2, \ldots, q$, and $i = 1, 2, \ldots, L$. Forming the log-likelihood function and eliminating irrelevant terms that are independent of l, we find that the maximization of the log-likelihood function is equivalent to selecting the largest of the q decision variables

$$U_l = \sum_{i=1}^{L}|y_{li}|^2, \quad l = 1,2,\ldots,q \qquad (2\text{-}103)$$

Consider noncoherent binary FSK. Because of the symmetry of the signals, $P_b(L)$ can be calculated by assuming that $s_1(t)$ was transmitted. Given that $s_1(t)$ was transmitted, the two decision variables at the combiner output are

$$\begin{aligned} U_1 &= \sum_{i=1}^{L}|2\mathcal{E}\alpha_i e^{j\theta_i}+N_{1i}|^2 \\ &= \sum_{i=1}^{L}\left(2\mathcal{E}\alpha_i\cos\theta_i+N_{1i}^R\right)^2+\sum_{i=1}^{L}\left(2\mathcal{E}\alpha_i\sin\theta_i+N_{1i}^I\right)^2 \qquad (2\text{-}104) \\ U_2 &= \sum_{i=1}^{L}|N_{2i}|^2=\sum_{i=1}^{L}\left(N_{2i}^R\right)^2+\sum_{i=1}^{L}\left(N_{2i}^I\right)^2 \qquad (2\text{-}105) \end{aligned}$$

where N_{1i} and N_{2i} are the independent, complex-valued, zero-mean, Gaussian noise variables defined by

$$N_{ki} = \int_0^T 2n_i(t)e^{-j2\pi f_c t}s_k^*(t)dt \ , \quad k = 1,2, \quad i = 1,2,\ldots,L \qquad (2\text{-}106)$$

and N_{ki}^R and N_{ki}^I are the real and imaginery parts of N_{ki}, respectively.

Since each $n_i(t)$ in (2-106) is a zero-mean, white Gaussian noise process with the same two-sided power spectral density $N_0/2$, (2-44), (2-106), and the spectral limitations of each $s_k(t)$ imply that

$$E[(N_{ki}^R)^2] = E[(N_{ki}^I)^2] = 2\mathcal{E}N_0 \ , \quad k = 1,2, \quad i = 1,2,\ldots,L \qquad (2\text{-}107)$$

and that N_{ki}^R and N_{ki}^I are uncorrelated, zero-mean, jointly Gaussian random variables and, hence, are independent of each other. Similarly, it can be verified by using the independence of $n_i(t)$ and $n_l(t)$, $i \neq l$, and the orthogonality of $s_1(t)$ and $s_2(t)$ that all $4L$ random variables in the sets $\{N_{ki}^R\}$ and $\{N_{ki}^I\}$ are statistically independent of each other. When independent, identically distributed, Rayleigh fading occurs in each branch, $\alpha_i \cos\theta_i$ and $\alpha_i \sin\theta_i$ are zero-mean, independent, Gaussian random variables with the same variance equal to $E[\alpha_i^2]/2 = E[\alpha_1^2]/2$, $i = 1, 2, \ldots, L$, as shown in Section A-4. Therefore, both U_1 and U_2 have central chi-square distributions with $2L$ degrees of freedom. From (A-18), the density function of U_k is

$$f_k(x) = \frac{1}{(2\sigma_k^2)^L(L-1)!}x^{L-1}\exp\left(-\frac{x}{2\sigma_k^2}\right)u(x), \quad k = 1,2 \qquad (2\text{-}108)$$

where (2-107) and (2-56) give

$$\sigma_2^2 = E[(N_{2i}^R)^2] = 2\mathcal{E}N_0 \qquad (2\text{-}109)$$

$$\sigma_1^2 = E[(2\mathcal{E}\alpha_1\cos\theta_i + N_{1i}^R)^2] = 2\mathcal{E}N_0(1+\bar{\gamma}) \qquad (2\text{-}110)$$

Since an erroneous decision is made if $U_2 > U_1$,

$$P_b(L) = \int_0^\infty \frac{x^{L-1}\exp\left(-\frac{x}{2\sigma_1^2}\right)}{(2\sigma_1^2)^L(L-1)!}\left[\int_x^\infty \frac{y^{L-1}\exp\left(-\frac{y}{2\sigma_2^2}\right)}{(2\sigma_2^2)^L(L-1)!}\,dy\right]dx \qquad (2\text{-}111)$$

Using (2-59) inside the brackets and integrating, we obtain

$$P_b(L) = \int_0^\infty \exp\left(-\frac{x}{2\sigma_1^2}\right)\sum_{i=0}^{L-1}\frac{(x/2\sigma_2^2)^i}{i!}\,\frac{x^{L-1}\exp\left(-\frac{x}{2\sigma_2^2}\right)}{(2\sigma_1^2)^L(L-1)!}\,dx \qquad (2\text{-}112)$$

Changing variables, applying (A-12), and simplifying gives (2-65), where the bit error probability for $L = 1$ is

$$p = \frac{1}{2+\bar{\gamma}} \qquad \text{(noncoherent FSK)} \qquad (2\text{-}113)$$

and $\bar{\gamma}$ is given by (2-56). Thus, $P_b(L)$ is once again given by (2-64). Equations (2-113) and (2-101) indicate that 3 dB more power is needed for noncoherent FSK to provide the same performance as DPSK. As discussed subsequently in Section 4.1, the performance of DPSK is approximately equaled by using minimum-shift keying and the configuration shown in Figure 12.

Equation (2-64) is valid for MRC and PSK or coherent FSK and also for EGC and DPSK or noncoherent FSK. Once the bit error probability in the absence of diversity combining, p, is determined, the bit error probability for diversity combining in the presence of independent Rayleigh fading, $P_b(L)$, can be calculated from (2-64). A plot of $P_b(L)$ versus p for different values of L is displayed in Figure 15. This figure illustrates the diminishing returns obtained as L increases. A plot of $P_b(L)$ versus $\bar{\gamma}$, the SNR per branch for one bit, is displayed in Figure 16 for MRC with PSK and EGC with DPSK and noncoherent FSK. The plot for MRC with coherent FSK is nearly the same as that for EGC with DPSK. Since (2-67) is valid for all these modulations, we find that $P_b(L)$ is asymptotically proportional to $\bar{\gamma}^{-L}$ with only the proportionality constant differing among the modulation types.

For noncoherent q-ary orthogonal signals such as MFSK, the union bound and the preceding analysis indicates that the symbol error probability is

$$P_s(L) \leq (q-1)\left\{p - (1-2p)\sum_{i=1}^{L}\binom{2i-1}{i}[p(1-p)]^i\right\} \tag{2-114}$$

where

$$p = \frac{1}{2 + (\log_2 q)\bar{\gamma}} \tag{2-115}$$

and the factor $\log_2 q$ accounts for the number of bits included in each nonbinary symbol. Relations (2-114) and (2-115) indicate that for fixed values of L, $P_s(L)$ decreases slowly as q increases. This result is confirmed by an exact analysis [3], [7].

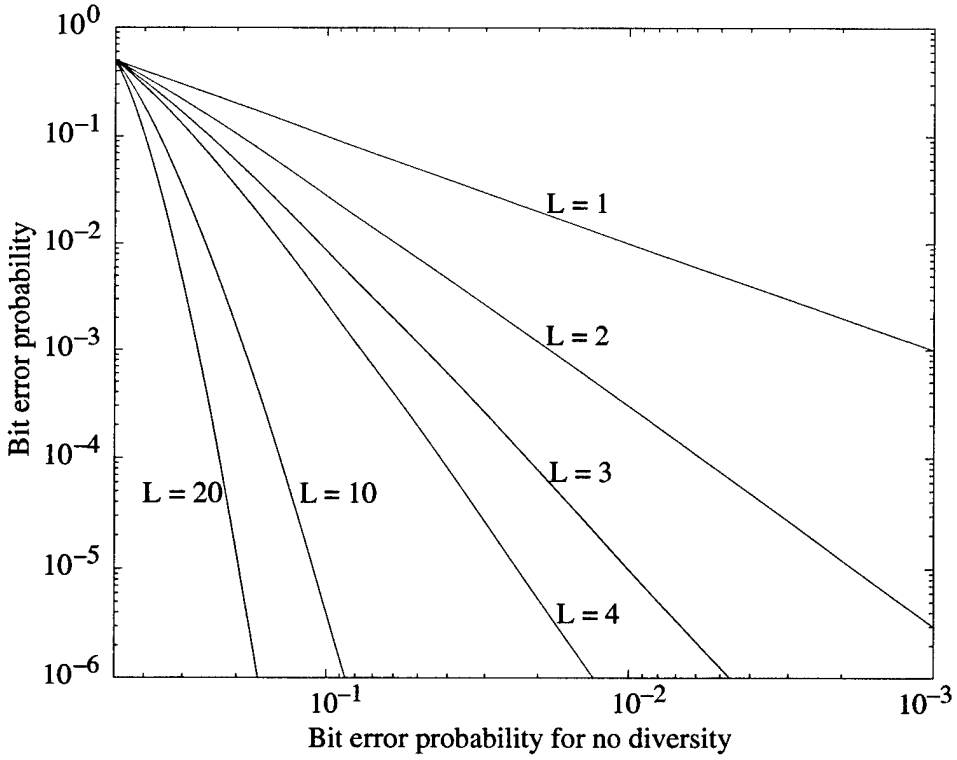


Figure 15. Bit error probability for MRC with PSK and coherent FSK and for EGC with DPSK and noncoherent FSK.

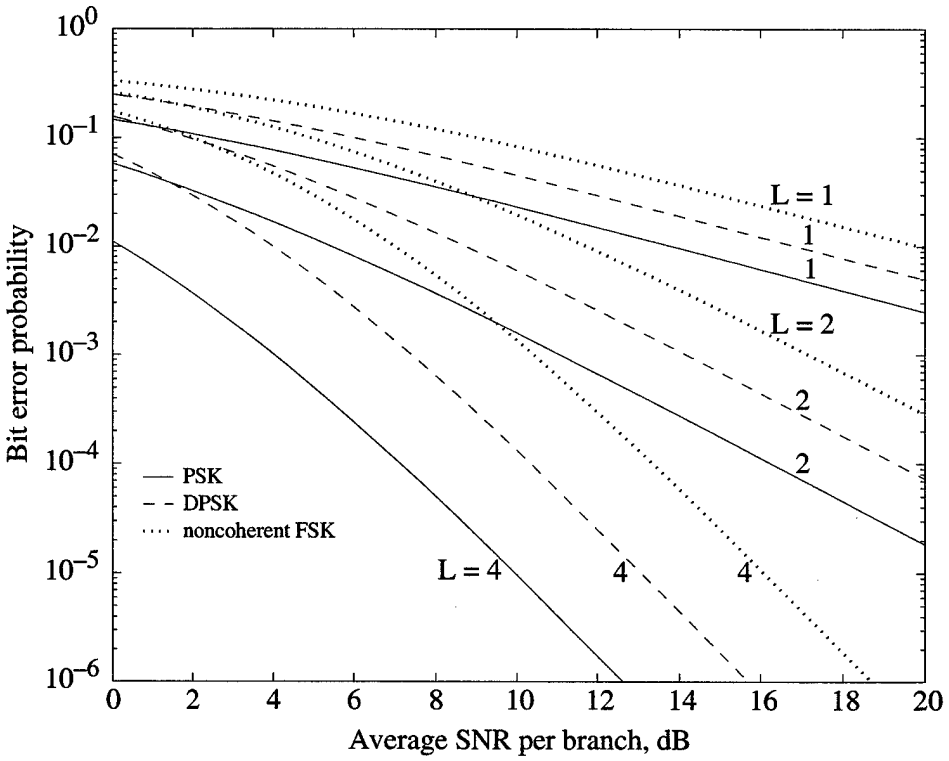


Figure 16. Bit error probability for MRC with PSK and for EGC with DPSK and noncoherent FSK.

2.5 Selection Diversity

A *selection-diversity system* selects the diversity branch that has the largest signal-to-noise ratio and forwards the signal in this branch for further processing. In a fading environment, selection diversity is sensible only if the selection rate is much faster than the fading rate. If the noise and interference levels in all the branches are nearly the same, then the total signal-plus-noise power in each branch rather than the signal-to-noise ratio can be measured to enable the selection process, thereby allowing a major simplification. Selection diversity does not provide a performance as good as maximal-ratio combining or equal-gain combining when the interference plus noise in each branch is uncorrelated with that in the other branches. However, selection diversity requires only a single demodulator, and when noises or interference signals are correlated, then selection diversity may become more competitive.

If the noise in each diversity branch is zero-mean and $E[|n_i|^2] = 2\sigma_i^2$, then the SNR in branch i is $\rho_i = p_s\alpha_i^2/\sigma_i^2$. If each of the $\{\alpha_i\}$ has a Rayleigh distribution and $\sigma_i^2 = \sigma^2$, $i = 1, 2, \ldots, L$, then the SNR in each branch has the same expected value

$$\bar{\rho} = \frac{p_s}{\sigma^2} E[\alpha_1^2] \tag{2-116}$$

The results of Appendix A-4 for the square of a Rayleigh-distributed random variable indicate that each SNR has the exponential probability density function

$$f_\rho(x) = \frac{1}{\bar{\rho}} \exp\left(-\frac{x}{\bar{\rho}}\right) u(x) \tag{2-117}$$

The corresponding probability distribution function is

$$F_\rho(x) = \left[1 - \exp\left(-\frac{x}{\bar{\rho}}\right)\right] u(x) \tag{2-118}$$

The branch with the largest SNR is selected. The probability that the SNR of the selected branch is less than or equal to x is equal to the probability that all the branch SNR's are simultaneously less than or equal to x. Therefore, the probability distribution function of the SNR of the selected branch is

$$F_{\rho 0}(x) = \left[1 - \exp\left(-\frac{x}{\bar{\rho}}\right)\right]^L u(x) \tag{2-119}$$

The corresponding probability density function is

$$f_{\rho 0}(x) = \frac{L}{\bar{\rho}} \exp\left(-\frac{x}{\bar{\rho}}\right) \left[1 - \exp\left(-\frac{x}{\bar{\rho}}\right)\right]^{L-1} u(x) \tag{2-120}$$

The average SNR obtained by selection diversity is calculated by integrating the SNR over the density given by (2-120). The result is

$$\begin{aligned} E[\rho_0] &= \int_0^\infty \frac{L}{\bar{\rho}} x \exp\left(-\frac{x}{\bar{\rho}}\right) \left[1 - \exp\left(-\frac{x}{\bar{\rho}}\right)\right]^{L-1} dx \\ &= \bar{\rho} L \int_0^\infty x e^{-x} \left(\sum_{i=0}^{L-1} \binom{L-1}{i} (-1)^i e^{-xi}\right) dx \\ &= \bar{\rho} \sum_{i=1}^{L} \binom{L}{i} \frac{(-1)^{i+1}}{i} \end{aligned} \tag{2-121}$$

The second equality results from a change of variable and the substitution of the binomial expansion. The third equality results from a term-by-term integration using (A-12) and an algebraic simplification. Substituting (2-116) and using a known series identity [8], we obtain

$$E[\rho_0] = \frac{p_s}{\sigma^2} E[\alpha_1^2] \sum_{i=1}^{L} \frac{1}{i} \tag{2-122}$$

Thus, the average SNR for selection diversity with $L \geq 2$ is less than that for MRC and the EGC, as indicated by (2-36) and (2-90), respectively. Approximating the summation in (2-122) by an integral, it is observed that the ratio of the average SNR for MRC to that for selection diversity is approximately $L/\ln L$ for $L \geq 2$.

Whether preprocessing or postprocessing is used, the performance of a receiver with selection diversity is the same. Suppose that the modulation

is PSK and optimal coherent demodulation follows the selection process. From (2-45), it follows that the conditional bit error probability is again given by the right-hand side of (2-52) with

$$\gamma_b = \frac{\mathcal{E}}{N_0} \max_i \left(\alpha_i^2\right) \tag{2-123}$$

If the $\{\alpha_i\}$ have identical Rayleigh distribution functions, then a derivation similar to the one leading to (2-120) indicates that the density function of γ_b is given by (2-120) with $\bar{\gamma}$ in place of $\bar{\rho}$, where $\bar{\gamma}$ is defined by (2-56). Therefore, using the binomial expansion, the bit error probability is

$$\begin{aligned} P_b(L) &= \int_0^\infty Q(\sqrt{2x}) \frac{L}{\bar{\gamma}} \exp\left(-\frac{x}{\bar{\gamma}}\right) \left[1 - \exp\left(-\frac{x}{\bar{\gamma}}\right)\right]^{L-1} dx \\ &= \sum_{i=0}^{L-1} \binom{L-1}{i} (-1)^i \frac{L}{\bar{\gamma}} \int_0^\infty Q(\sqrt{2x}) \exp\left[-x\left(\frac{1+i}{\bar{\gamma}}\right)\right] dx \end{aligned} \tag{2-124}$$

The last integral can be evaluated in the same manner as the one in (2-57). After regrouping factors, the result is

$$P_b(L) = \frac{1}{2} \sum_{i=0}^{L-1} \binom{L}{i+1} (-1)^i \left(1 - \sqrt{\frac{\bar{\gamma}}{i+1+\bar{\gamma}}}\right) \qquad \text{(PSK, QPSK)} \tag{2-125}$$

This equation is valid for QPSK since it can be implemented as two parallel binary PSK waveforms.

For coherent FSK, the conditional bit error probability is $P_b(\gamma_b) = Q(\sqrt{\gamma_b})$. Therefore, it is found that

$$P_b(L) = \frac{1}{2} \sum_{i=0}^{L-1} \binom{L}{i+1} (-1)^i \left(1 - \sqrt{\frac{\bar{\gamma}}{2i+2+\bar{\gamma}}}\right) \qquad \text{(coherent FSK)} \tag{2-126}$$

Again, 3 dB more power is needed to provide to the same performance as PSK.

When DPSK is the data modulation, the conditional bit error probability is $\exp(-\gamma_b)/2$. Thus, selection diversity provides the bit error probability

$$P_b(L) = \int_0^\infty \frac{L}{2\bar{\gamma}} \exp\left[\left(-x\frac{1+\bar{\gamma}}{\bar{\gamma}}\right)\right] \left[1 - \exp\left(-\frac{x}{\bar{\gamma}}\right)\right]^{L-1} dx \tag{2-127}$$

The *beta function* is defined as

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt \;, \quad x > 0, \quad y > 0 \tag{2-128}$$

If y is a positive integer n, then the substitution of the binomial expansion of $(1-t)^{n-1}$ and the evaluation of the resulting integral yields

$$B(x,n) = \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{(-1)^i}{i+x} \;, \quad n \geq 1 \;, \quad x > 0 \tag{2-129}$$

Using $t = \exp(-x/\bar{\gamma})$ to change the integration variable in (2-127) and then using (2-128) gives

$$P_b(L) = \frac{L}{2} B(\bar{\gamma} + 1, L) \qquad \text{(DPSK)} \tag{2-130}$$

For noncoherent MFSK, the classical formula for the conditional symbol error probability given the $\{\alpha_i\}$ is [3], [8]

$$P_{s|\boldsymbol{\alpha}}(\gamma_b) = \sum_{i=1}^{q-1} \frac{(-1)^{i+1}}{i+1} \binom{q-1}{i} \exp\left(-\frac{i\gamma_b}{i+1}\right) \tag{2-131}$$

Therefore, a derivation similar to that of (2-130) yields the symbol error probability

$$P_s(L) = L \sum_{i=1}^{q-1} \frac{(-1)^{i+1}}{i+1} \binom{q-1}{i} B\left(1 + \frac{i\bar{\gamma}}{i+1}, L\right) \qquad \text{(noncoherent MFSK)} \tag{2-132}$$

For binary FSK, the bit error probability is

$$P_b(L) = \frac{L}{2} B\left(\frac{\bar{\gamma}}{2} + 1, L\right) \qquad \text{(noncoherent FSK)} \tag{2-133}$$

which exhibits the usual 3 dB disadvantage compared with DPSK.

Asymptotic forms of (2-130) and (2-133) may be obtained by substituting

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{2-134}$$

To prove this identity, let $y = z^2$ in the integrand of the gamma function defined in (A-12). Express the product $\Gamma(a)\Gamma(b)$ as a double integral, change to polar coordinates, integrate over the radius to obtain a result proportional to $\Gamma(a+b)$, and then change the variable in the remaining integral to obtain $B(a,b)\Gamma(a+b)$.

For DPSK, the substitution of (2-134) and (2-101) into (2-130) and the use of $\Gamma(\bar{\gamma}+L+1) = (\bar{\gamma}+L)(\bar{\gamma}+L-1)\ldots(\bar{\gamma}+1)\Gamma(\bar{\gamma}+1) \geq (\bar{\gamma}+1)^L\Gamma(\bar{\gamma}+1)$ give

$$P_b(L) \leq 2^{L-1} L\,!\, p^L \tag{2-135}$$

For noncoherent FSK, a similar derivation using (2-113) and (2-133) yields the same upper bound, which is tight when $\bar{\gamma} >> L$. The upper bound on $P_b(L)$ for DPSK and noncoherent FSK with EGC is given by (2-67). Comparing the latter with (2-135) indicates the disadvantage of selection diversity relative to EGC when $\bar{\gamma} >> L$ and $L \geq 2$.

Figure 17 shows $P_b(L)$ as a function of the average SNR per branch, assuming selection diversity with PSK, DPSK, and noncoherent FSK. A comparison of Figures 17 and 16 indicates the reduced gain provided by selection diversity relative to MRC and EGC.

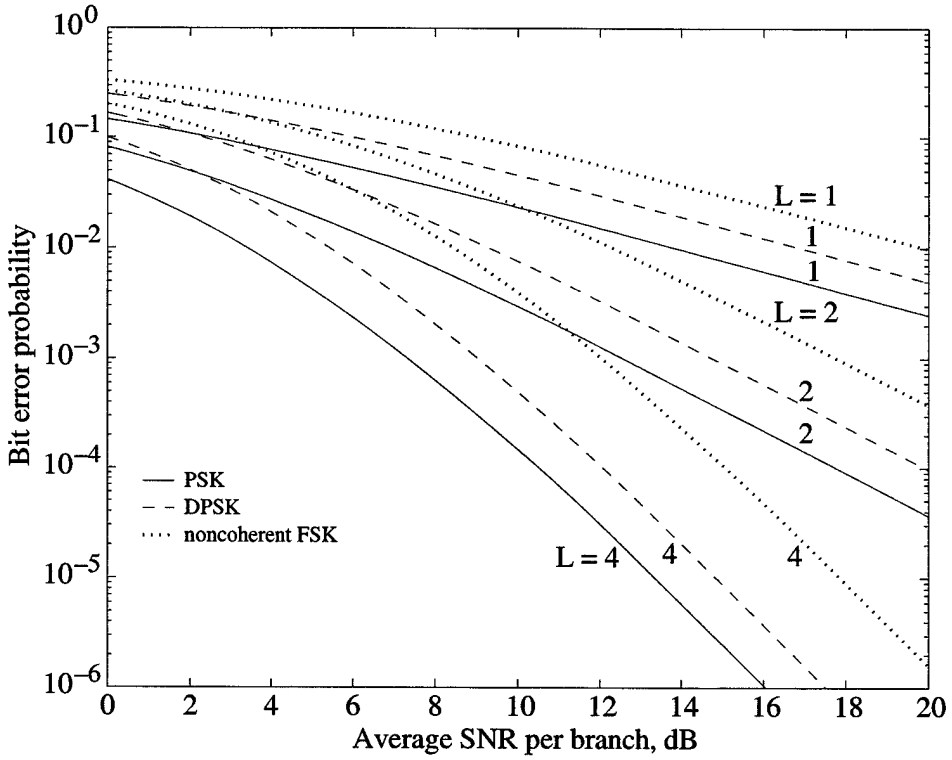


Figure 17. Bit error probability for selection diversity with PSK, DPSK, and noncoherent FSK.

A fundamental limitation of selection diversity is made evident by the plane-wave example in which the signal and interference steering vectors are given by (2-91) and (2-92). In this example, the SNR's are equal in all the diversity branches. Consequently, selection diversity can give no better performance than no diversity combining or the use of a single branch. In contrast, (2-98) indicates that EGC can improve the SINR significantly.

2.6 Rake Receiver

A rake receiver provides *path diversity* by coherently combining resolvable multipath components that are often present during frequency-selective fading. Consider a multipath channel with frequency-selective fading slow enough that its time variations are negligible over a signaling interval. To harness the energy in all the multipath components, a receiver should decide which signal was transmitted among M candidates, $s_1(t)$, $s_2(t)$, $\ldots$, $s_M(t)$, only after processing all the received multipath components of the signal. Thus, the receiver selects among the M baseband signals or complex envelopes

$$v_k(t) = \sum_{i=1}^{L} c_i s_k(t - \tau_i), \quad k = 1, 2, \ldots, M, \quad 0 \le t \le T + T_d \qquad (2\text{-}136)$$

where T is the duration of the transmitted signal, T_d is the multipath delay spread, L is the number of multipath components, τ_i is the delay of component i, and the channel parameter c_i is a complex number representing the attenuation and phase shift of component i. An idealized sketch

of the output of a baseband matched filter that receives three multipath components of the signal to which it is matched is shown in Figure 18. If a signal has bandwidth W, then the duration of the matched-filter response to this signal is on the order of $1/W$. Multipath components that produce distinguishable matched-filter output pulses are said to be *resolvable*. Thus, three multipath components are resolvable if their relative delays are greater than $1/W$, as depicted in the figure. A necessary condition for at least two resolvable multipath components is that duration $1/W$ is less than the delay spread T_d. From (1-52) it follows that $W > B_c$ is required, which implies that frequency-selective fading and resolvable multipath components are associated with wideband signals. There are at most $\lfloor T_d W \rfloor + 1$ resolvable components, where $\lfloor x \rfloor$ denotes the largest integer in x. As observed in the figure, intersymbol interference at the sampling times is not significant if $T_d + 1/W$ is less than the symbol duration T_s.

For the following analysis, it is assumed that the M possible signals are orthogonal to each other and that the data symbols are independent of each other so that the maximum-likelihood receiver makes symbol-by-symbol decisions [3], [6], [10]. This receiver uses a separate baseband matched filter or correlator for each possible desired signal including its multipath components. Thus, if $s_k(t)$ is the kth symbol waveform, $k = 1, 2, \ldots, M$, then the kth matched filter is matched to the signal $v_k(t)$ in (2-136) with $T = T_s$. Each matched-filter output sampled at $t = T_s + T_d$ provides a decision variable. A derivation similar to that of (2-74) indicates that the kth decision variable is

$$U_k = \text{Re}\left[\sum_{i=1}^{L} c_i^* \int_0^{T_s+T_d} r(\tau) s_k^*(\tau - \tau_i) d\tau\right] \tag{2-137}$$

where $r(t)$ is the received signal, including the noise, after translation to baseband. A receiver implementation based on this equation would require a separate transversal filter or delay line and a matched filter for each possible waveform $s_k(t)$. An alternative form that requires only a single transversal filter and M matched filters is derived by changing variables in (2-137) and using the fact that $s_k(t)$ is zero outside the interval $[0, T_s)$. The result is

$$U_k = \text{Re}\left[\sum_{i=1}^{L} c_i^* \int_0^{T_s} r(\tau + \tau_i) s_k^*(\tau) d\tau\right] \tag{2-138}$$

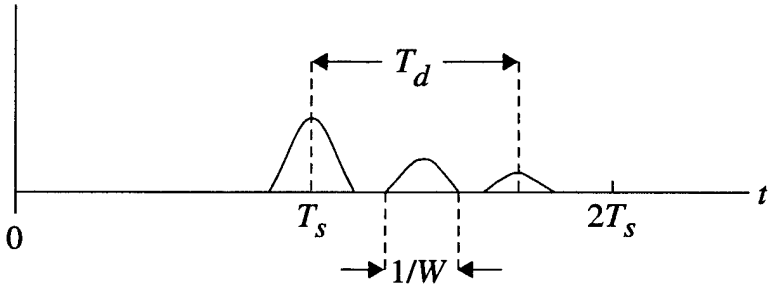


Figure 18. Response of matched filter to input with three resolvable multipath components.

For frequency-selective fading and resolvable multipath components, a simplifying assumption is that each delay is an integer multiple of $1/W$. Accordingly, L is increased to equal the maximum number of resolvable components, and we set $\tau_i = (i-1)/W$, $i = 1, 2, \ldots, L$, and $(L-1)/W \approx \tau_m$, where τ_m is the maximum delay. As a result, some of the $\{c_i\}$ may be equal to zero. The decision variables become

$$U_k = \text{Re}\left[\sum_{i=1}^{L} c_i^* \int_0^{T_s} r(\tau + (i-1)/W) s_k^*(\tau) d\tau\right], \quad k = 1, 2, \ldots, M \quad \text{(2-139)}$$

A receiver based on these decision variables, which is called a *rake receiver*, is diagrammed in Figure 19. Since $r(t)$ is designated as the output of the final tap, the sampling occurs at $t = T_s$. Each tap output contains at most one multipath component of $r(t)$.

The rake receiver requires that the channel parameters $\{c_i\}$ be known or estimated. An estimation might be done by applying each tap output to M parallel matched filters after a one-symbol delay. The previous symbol decision is used to select one matched-filter output for each tap output. The L matched-filter outputs are lowpass-filtered to provide estimates of the channel parameters. The estimates must be updated at a rate exceeding the fade rate of (1-54) or (1-56).

An alternative configuration is to use a separate transversal filter for each decision variable and to move the corresponding matched filter to the front, as shown in Figure 20(a). The matched-filter or correlator output is applied to L_s parallel *fingers*, the outputs of which are recombined and sampled to produce the decision variable. The number of fingers L_s, where $L_s \leq L$, is equal to the number the resolvable components that have significant power. The matched filter produces a number of output pulses in response to the multipath components, as illustrated in Figure 18. Each finger delays and weights one of these pulses by the appropriate amount so that all the finger output pulses are aligned in time and can be constructively combined after weighting, as shown in Figure 20(b). Digital devices can be used because the sampling immediately follows the matched filtering. The delay of each significant multipath component may be estimated by using envelope detectors and threshold devices. Let t_e denote the time required to estimate the relative delay of a multipath component, and let v denote

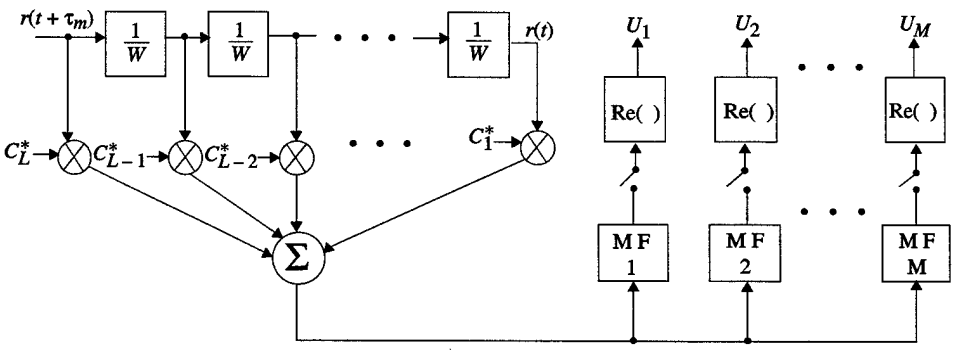


Figure 19. Rake receiver for M orthogonal pulses. MF denotes a matched filter.

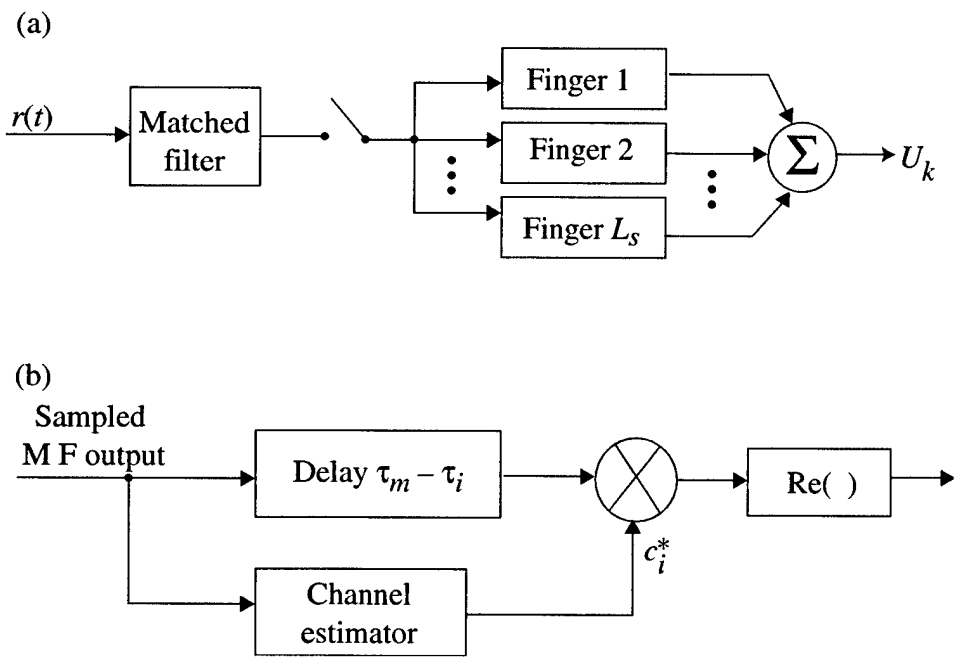


Figure 20. Rake receiver: (a) basic configuration for generating a decision variable and (b) a single finger.

the relative radial velocity of a receiver relative to a transmitter. Then vt_e/c is the change in delay that occurs during the estimation procedure, where c is the speed of an electromagnetic wave. This change must be much less than the duration of a multipath output pulse shown in Figure 18 if the delay estimate is to be useful. Thus, with v interpreted as the maximum speed of a mobile in a mobile communications network,

$$t_e << \frac{c}{vW} \tag{2-140}$$

is required of the multipath-delay estimation.

Suppose that $s_k(t)$ is a direct-sequence signal with chip duration $T_c = 1/W$. If the processing gain T_s/T_c is large, the spreading sequence has a small autocorrelation when the relative delay is T_c or more, and

$$\int_0^{T_s} s_k(t + (i-1)/W) s_k(t) dt << \int_0^{T_s} |s_k(t)|^2 dt \ , \ i \geq 2 \tag{2-141}$$

When the data modulation is binary antipodal or PSK, only a single symbol waveform $s_1(t)$ and its associated decision variable U_1 are needed. After translation to baseband, the received signal is

$$r(t) = \left\{ \mathrm{Re}\left[v_1(t) e^{j2\pi f_c t} \right] + n(t) \right\} 2e^{-j2\pi f_c t} \tag{2-142}$$

where $v_1(t)$ is given by (2-136) and $n(t)$ is zero-mean white Gaussian noise. Let $\alpha_i = |c_i|$ and assume that α_1 is larger than or comparable to α_i, $i \neq 1$. Substituting (2-142) and (2-136) into (2-139) with $k = 1$ and then using (2-141), we again obtain (2-47). Thus, the rake receiver produces MRC, and the conditional bit error probability given the $\{\alpha_i\}$ is provided by (2-52).

However, for a rake receiver, each of the $\{\alpha_i\}$ is associated with a different multipath component, and hence each $E[\alpha_i]$ has a different value in general. Therefore, the derivation of $P_b(L)$ must be modified.

Equation (2-53) may be expressed as

$$\gamma_b = \sum_{i=1}^{L} \gamma_i \quad , \quad \gamma_i = \frac{\mathcal{E}}{N_0}\alpha_i^2 \tag{2-143}$$

If each α_i has a Rayleigh distribution then each γ_i has the exponential probability density function (Appendix A-4)

$$f_{\gamma_i}(x) = \frac{1}{\bar{\gamma}_i} \exp\left(-\frac{x}{\bar{\gamma}_i}\right) u(x), \quad i = 1, 2, \ldots, L \tag{2-144}$$

where the average SNR for a bit in branch i is

$$\bar{\gamma}_i = \frac{\mathcal{E}}{N_0} E[\alpha_i^2] \ , \ i = 1, 2, \ldots, L \tag{2-145}$$

If each multipath component fades independently so that each of the $\{\gamma_i\}$ is statistically independent, then γ_b is the sum of independent, exponentially distributed random variables. The results of Appendix A-5 indicate that the probability density function of γ_b is

$$f_{\gamma_b}(x) = \sum_{i=1}^{L} \frac{A_i}{\bar{\gamma}_i} \exp\left(-\frac{x}{\bar{\gamma}_i}\right) u(x) \tag{2-146}$$

where

$$A_i = \begin{cases} \prod\limits_{\substack{k=1 \\ k \neq i}}^{L} \dfrac{\bar{\gamma}_i}{\bar{\gamma}_i - \bar{\gamma}_k} & , \quad L \geq 2 \\ 1 \, , & L = 1 \end{cases} \tag{2-147}$$

The bit error probability is determined by averaging the conditional bit error probability $P_2(\gamma_b) = Q(\sqrt{2\gamma_b})$ over the density given by (2-146). A derivation similar to that leading to (2-61) yields

$$P_b(L) = \frac{1}{2} \sum_{i=1}^{L} A_i \left(1 - \sqrt{\frac{\bar{\gamma}_i}{1 + \bar{\gamma}_i}}\right) \qquad \text{(PSK, QPSK)} \tag{2-148}$$

The number of fingers in an ideal rake receiver equals the number of significant resolvable multipath components, which is constantly changing in a mobile communications receiver. Rather than attempting to implement all the required fingers that may sometimes be required, a more practical alternative is to implement a fixed number of fingers independent of the number of multipath components. *Generalized selection diversity* entails selecting the L_c strongest resolvable components among the L available ones

and then applying MRC or EGC of these L_c components, thereby discarding the $L - L_c$ components with the lowest SNRs. Analysis [2] indicates that diminishing returns are obtained as L_c increases, but for a fixed value of L_c, the performance improves as L increases.

An increase in the number of resolved components L is potentially beneficial if it is caused by natural changes in the physical environment that generate additional multipath components. However, an increase in L due to an increase in the bandwidth W is not always beneficial [11]. Although new components provide additional diversity and may exhibit the more favorable Ricean fading rather than Rayleigh fading, the average power per multipath component decreases because some composite components fragment into more numerous but weaker components. Hence, the estimation of the channel parameters becomes more difficult, and the fading of some multipath components may be highly correlated rather than independent.

The estimation of the channel parameters needed in a rake receiver becomes more difficult as the fading rate increases. When the estimation errors are large, it may be preferable to use a rake receiver that avoids channel-parameter estimation by abandoning MRC and using noncoherent postdetection EGC. The form of this rake receiver for binary signals is depicted in Figure 21. Each tap output of the transversal filter provides an input to the equal-gain combiner, which may have the form of Figure 13 or Figure 14.

For two orthogonal signals that satisfy (2-141) and the rake receiver of Figures 21 and 14, the decision variables are given by (2-104) and (2-105). Since U_2 has a central chi-square distribution with $2L$ degrees of freedom, the probability density function of U_2 is given by (2-108) and (2-109). Equation (2-104) can be expressed as

$$U_1 = \sum_{i=1}^{L} \left[\left(2\mathcal{E}\alpha_i \cos\theta_i + N_{1i}^R \right)^2 + \left(2\mathcal{E}\alpha_i \sin\theta_i + N_{1i}^I \right)^2 \right] \qquad (2\text{-}149)$$

Each phase θ_i is assumed to be statistically independent and uniformly distributed over $[0, 2\pi)$. Since each α_i has a Rayleigh distribution, $\alpha_i \cos\theta_i$ and

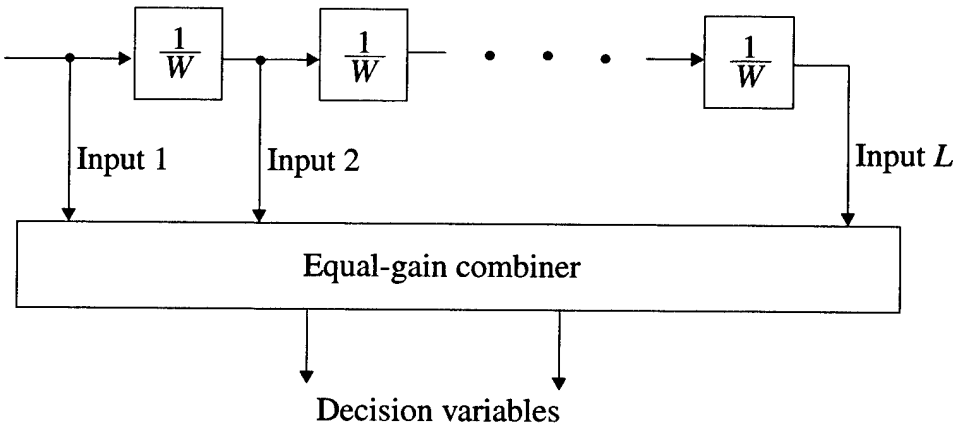


Figure 21. Rake receiver that uses equal-gain combiner to avoid channel-parameter estimation.

$\alpha_i \sin\theta_i$ have zero-mean, independent, Gaussian distributions. Therefore, as indicated in Appendix A-4, each term of U_1 has an exponential distribution with mean

$$m_i = 4\mathcal{E}N_0(1+\bar{\gamma}_i) \tag{2-150}$$

where $\bar{\gamma}_i$ is defined by (2-145). Since the statistical independence of the $\{\alpha_i\}$ and $\{\theta_i\}$ implies the statistical independence of the terms of U_1, the probability density function of U_1 for distinct values of the $\{\bar{\gamma}_i\}$ is given by (A-45) and (A-46) with $N = L$. Since an erroneous decision is made if $U_2 > U_1$,

$$P_b(L) = \sum_{i=1}^{L} \frac{B_i}{m_i} \int_0^\infty \exp\left(-\frac{x}{m_i}\right) \int_x^\infty \frac{y^{L-1}\exp\left(-\frac{y}{2\sigma_2^2}\right)}{(2\sigma_2^2)^L(L-1)!}dy\,dx \tag{2-151}$$

Integrating by parts to eliminate the inner integral, changing the remaining integration variable, applying (A-12), and simplifying yields the bit error probability for orthogonal signals and a rake receiver with noncoherent postdetection EGC:

$$P_b(L) = \sum_{i=1}^{L} B_i \left[1 - \left(\frac{1+\bar{\gamma}_i}{2+\bar{\gamma}_i}\right)^L\right] \quad \text{(orthogonal signals)} \tag{2-152}$$

where

$$B_i = \begin{cases} \prod\limits_{\substack{k=1 \\ k\neq i}}^{L} \dfrac{1+\bar{\gamma}_i}{\bar{\gamma}_i - \bar{\gamma}_k}, & L \geq 2 \\ 1, & L = 1 \end{cases} \tag{2-153}$$

An alternative derivation of (2-152) using the direct-conversion receiver modeled in Appendix B-3 is given in [12]. Equation (2-152) is more compact and considerably easier to evaluate than the classical formula of Proakis [3], which is derived in a different way.

Another way to avoid channel-parameter estimation is to use DPSK and the diversity receiver of Figure 13 in Figure 21. The classical analysis [3] implies that $P_b(L)$ is given by (2-152) and (2-153) with $\bar{\gamma}_i$ replaced by $2\bar{\gamma}_i$.

For dual rake combining with orthogonal signals, (2-152) reduces to

$$P_b(2) = \frac{8 + 5\bar{\gamma}_1 + 5\bar{\gamma}_2 + 3\bar{\gamma}_1\bar{\gamma}_2}{(2+\bar{\gamma}_1)^2(2+\bar{\gamma}_2)^2} \tag{2-154}$$

If $\bar{\gamma}_2 = 0$, then

$$P_b(2) = \frac{2 + \frac{5}{4}\bar{\gamma}_1}{(2+\bar{\gamma}_1)^2} \geq \frac{1}{2+\bar{\gamma}_1} = P_b(1) \tag{2-155}$$

This result illustrates the performance degradation that results when a rake combiner uses an input that provides no desired-signal component, which

may occur when EGC is used rather than MRC. In the absence of a desired-signal component, this input contributes only noise to the combiner.

If an adaptive array produces a directional beam to reject interference or enhance the desired signal, it also reduces the delay spread of the multipath components of the desired signal because components arriving from angles outside the beam are greatly attenuated. As a result, the potential benefit of a rake receiver diminishes. Another procedure is to assign a separate set of adaptive weights to each significant multipath component. Consequently, the adaptive array can form separate array patterns, each of which enhances a particular multipath component while nulling other components. The set of enhanced components are then applied to the rake receiver [13].

2.7 Error-Correcting Codes

If the channel symbols are interleaved to a depth beyond the coherence time of the channel, then the symbols fade independently. As a result, an error-correcting code provides a form of time diversity. Consider an (n, k) linear block code with soft-decision decoding, where n is the number of code symbols and k is the number of information symbols. Each possible codeword is assigned a number called the *metric*, which is a function of both the codeword and the demodulator output samples. A soft-decision decoder selects the codeword with the largest metric. The information bits are then recovered from this codeword.

Let $\mathbf{y}$ denote the n-dimensional vector of noisy output samples $y_i, i = 1, 2, \ldots, n$, produced by a demodulator that receives a sequence of n symbols. Let $\mathbf{x}_l$ denote the lth codeword vector with symbols $x_{li}, i = 1, 2, \ldots, n$. Let $p(\mathbf{y}|\mathbf{x}_l)$ denote the *likelihood function*, which is the conditional probability density function of $\mathbf{y}$ given that $\mathbf{x}_l$ was transmitted. Let q denote the alphabet size of the code symbols, and let k denote the number of information symbols per codeword. The *maximum-likelihood decoder* finds the value of l, $1 \leq l \leq q^k$, for which the likelihood function is largest. If this value is l_0, the decoder decides that codeword l_0 was transmitted. Any monotonically increasing function of $p(\mathbf{y}|\mathbf{x}_l)$ can serve as the metric of a maximum-likelihood decoder. If $p(\mathbf{y}|\mathbf{x}_l)$ is known, the decoder *metric* is usually chosen to be a linear function of the logarithm of $p(\mathbf{y}|\mathbf{x}_l)$, which is called the *log-likelihood function*. If the demodulator outputs are statistically independent and a single output corresponds to each code symbol, then the likelihood function is the product of n conditional probability density functions, and the log-likelihood function *or maximum-likelihood metric* for each of the q^k possible codewords is

$$\ln [f(\mathbf{y}|\mathbf{x}_l)] = \sum_{i=1}^{n} \ln [f(y_i|x_{li})] , \quad l = 1, 2, \ldots, q^k \tag{2-156}$$

where $f(y_i|x_{li})$ is the conditional probability density function of y_i given the value of x_{li}. In the subsequent analysis, it is always assumed that per-

fect symbol interleaving or sufficiently fast fading ensures the statistical independence of the demodulator outputs so that (2-156) is applicable.

For binary PSK over a fading channel in which the fading is constant over a symbol interval, the received signal representing symbol i of codeword l is

$$r_i(t) = \mathrm{Re}[\alpha_i e^{j\theta_i} x_{li} \psi(t) e^{j2\pi f_c t}] + n_i(t) \ , \quad i = 1, 2, \ldots, n \tag{2-157}$$

where α_i is a random variable that includes the effects of the fading, $x_{li} = +1$ when binary symbol i is a 1 and $x_{li} = -1$ when binary symbol i is a 0, and $\psi(t)$ is the symbol waveform. The noise process $n_i(t)$ is independent, zero-mean, white Gaussian noise with autocorrelation given by (2-44). When codeword l is received in the presence of white Gaussian noise, the translation to baseband and the matched-filter or correlator demodulator, which is matched to $\psi(t)$, produces the samples

$$\begin{aligned} y_i &= \int_0^{T_s} 2 r_i(t) e^{-j2\pi f_c t} \psi^*(t) dt \\ &= 2\mathcal{E}_s \alpha_i e^{j\theta_i} x_{li} + \int_0^{T_s} 2 n_i(t) e^{-j2\pi f_c t} \psi^*(t) dt \ , \quad i = 1, 2, \ldots, n \end{aligned} \tag{2-158}$$

where T_s denotes the symbol duration and the symbol energy is

$$\mathcal{E}_s = \frac{1}{2} \int_0^{T_s} |\psi(t)|^2 dt \tag{2-159}$$

Since $\psi(t)$ is the sole basis function for the signal space, these samples provide sufficient statistics; that is, they contain all the relevant information in the received signal [3], [6], [10].

The spectrum of $\psi(t)$ is assumed to be confined to $|f| < f_c$. Using this assumption, (2-159), and (2-44), we find that the Gaussian noise term in (2-158) has independent real and imaginary components with the same variance $2\mathcal{E}_s N_0$. Therefore, the conditional probability density function of y_i given the values of x_{li}, α_i, and θ_i is,

$$f(y_i | x_{li}, \alpha_i, \theta_i) = \frac{1}{4\pi \mathcal{E}_s N_0} \exp\left[- \frac{|y_i - 2\mathcal{E}_s \alpha_i e^{j\theta_i} x_{li}|^2}{4\mathcal{E}_s N_0} \right] , \quad i = 1, 2, \ldots, n \ , \quad l = 1, 2, \ldots, 2^k \tag{2-160}$$

Substituting this equation into (2-156) and then eliminating irrelevant terms and factors that do not depend on the codeword l, we obtain the *maximum-likelihood metrics for PSK*:

$$U(l) = \sum_{i=1}^{n} \mathrm{Re}\left[y_i \alpha_i e^{-j\theta_i} x_{li} \right] , \quad l = 1, 2, \ldots, 2^k \tag{2-161}$$

which serve as decision variables.

The error probability of a linear binary block (n, k) code can be derived by assuming that the all-zero codeword represented by $x_{1i} = -1, i = 1, 2, \ldots, n$, was transmitted. If codeword l has Hamming weight w, then the comparison of the metrics $U(1)$ and $U(l)$ depends only on the w terms in (2-161) that differ. The probability of an error is deciding between two codewords is equal to the probability that the $U(1) < U(l)$. Therefore, if each of the $\{\alpha_i\}$ is independent with the identical Rayleigh distribution and $E[\alpha_i^2] = E[\alpha_1^2]$, $i = 1, 2, \ldots, n$, a derivation similar to the one leading to (2-64) indicates that the two-codeword error probability is

$$P_2(w) = p - (1 - 2p) \sum_{i=1}^{w-1} \binom{2i-1}{i} [p(1-p)]^i \tag{2-162}$$

where

$$p = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}_s}{1 + \bar{\gamma}_s}}\right) \quad \text{(PSK, QPSK)} \tag{2-163}$$

the average SNR per binary code symbol is

$$\bar{\gamma}_s = \frac{\mathcal{E}_s}{N_0} E[\alpha_1^2] = \frac{r\mathcal{E}_b}{N_0} E[\alpha_1^2] = r\bar{\gamma}_b \quad \text{(binary symbols)} \tag{2-164}$$

$\mathcal{E}_b$ is the information-bit energy, r is the code rate, and $\bar{\gamma}_b$ is the average SNR per bit. The same equations are valid for both PSK and QPSK because the latter can be transmitted as two independent binary PSK waveforms in phase quadrature.

Successive applications of the union bound indicates that the word error probability for soft-decision decoding of block codes satisfies [3], [10]

$$P_w \leq \sum_{l=d_m}^{n} A_l P_2(l) \leq (q^k - 1) P_2(d_m) \tag{2-165}$$

where d_m is the minimum distance between codewords, q is the alphabet size, and A_l is the number of codewords with weight l. For cyclic block codes, the average Hamming weight of the information symbols in codewords of weight l is kl/n [14]. Thus, when an incorrect codeword of weight l is chosen instead of the all-zero codeword, the information-symbol error probability is l/n, and an upper bound on the information-symbol error probability for soft-decision decoding is given by

$$P_{is} \leq \sum_{l=d_m}^{n} \frac{l}{n} A_l P_2(l) \tag{2-166}$$

An incorrectly decoded information symbol is assumed to be equally likely to be any of the $(q-1)$ other symbols. Among those symbols, a given bit is incorrect in $q/2$ instances. Therefore, the information-bit error probability is

$$P_{ib} = \frac{q}{2(q-1)} P_{is} \tag{2-167}$$

For q-ary orthogonal symbol waveforms $s_1(t)$, $s_2(t)$, ..., $s_q(t)$, q matched filters are needed. The observation vector is $\mathbf{y} = [\mathbf{y}_1\ \mathbf{y}_2 \ldots \mathbf{y}_q]$, where each $\mathbf{y}_k$ is an n-dimensional row vector of output samples y_{ki}, $i = 1, 2, \ldots, n$, from matched-filter k, which is matched to $s_k(t)$. Suppose that symbol i of codeword l uses $s_\nu(t)$. Because the symbol waveforms are orthogonal, when codeword l is received in the presence of white Gaussian noise, matched-filter k produces the samples

$$y_{ki} = 2\mathcal{E}_s \alpha_i e^{j\theta_i} \delta_{k\nu} + \int_0^{T_s} 2n_i(t) e^{-j2\pi f_c t} s_k^*(t) dt\,, \quad i = 1, 2, \ldots, n\,, \quad k = 1, 2, \ldots, q \qquad (2\text{-}168)$$

where $\delta_{k\nu} = 1$ if $k = \nu$ and $\delta_{k\nu} = 0$ otherwise, and the symbol energy for all the waveforms is

$$\mathcal{E}_s = \frac{1}{2} \int_0^{T_s} |s_k(t)|^2 dt\,, \quad k = 1, 2, \ldots, q \qquad (2\text{-}169)$$

Since each symbol waveform represents $\log_2 q$ bits, the average SNR per code symbol is

$$\bar{\gamma}_s = (\log_2 q) r \bar{\gamma}_b \qquad (2\text{-}170)$$

which reduces to (2-164) when $q = 2$.

If the spectra of the $\{s_k(t)\}$ are confined to $|f| < f_c$, then (2-169) and (2-44) imply that the real and imaginary components of the Gaussian noise term in (2-168) are independent and have the same variance $2\mathcal{E}_s N_0$. Therefore, the conditional probability density function of y_{ki} given the values of l, α_i, and θ_i is

$$f(y_{ki} | l, \alpha_i, \theta_i) = \frac{1}{4\pi \mathcal{E}_s N_0} \exp\left[- \frac{|y_{ki} - 2\mathcal{E}_s \alpha_i e^{j\theta_i} \delta_{k\nu}|^2}{4\mathcal{E}_s N_0} \right] \qquad (2\text{-}171)$$

The orthogonality of the $\{s_k(t)\}$, the independence of the white noise from symbol to symbol, and (2-44) imply the conditional independence of the $\{y_{ki}\}$.

For coherent MFSK, the $\{\alpha_i\}$ and the $\{\theta_i\}$ are assumed to be known, and the likelihood function is the product of qn densities given by (2-171) for $k = 1, 2, \ldots, q$ and $i = 1, 2, \ldots, n$. Forming the log-likelihood function and eliminating irrelevant terms that are independent of l, we obtain the *maximum-likelihood metrics for coherent MFSK*:

$$U(l) = \sum_{i=1}^{n} \mathrm{Re}\left[\alpha_i e^{-j\theta_i} V_{li} \right], \quad l = 1, 2, \ldots, q^k \qquad (2\text{-}172)$$

where $V_{li} = y_{\nu i}$ is the sampled output of the filter matched to $s_\nu(t)$, the signal representing symbol i of codeword l.

For a linear block code, the error probabilities may be calculated by assuming that the all-zero codeword denoted by $l = 1$ was transmitted. The comparison of the metrics $U(1)$ and $U(l)$, $l \neq 1$, depends only on the w terms

that differ, where w is the weight of codeword l. The two-codeword error probability is equal to the probability that $U(1) < U(l)$. For coherent MFSK and independent, identically distributed Rayleigh fading of each codeword symbol, a derivation similar to the one leading to (2-162) indicates that the two-codeword error probability is again given by $P_2(w)$ provided that

$$p = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}_s}{2 + \bar{\gamma}_s}}\right) \quad \text{(coherent MFSK)} \tag{2-173}$$

where $\bar{\gamma}_s$ is given by (2-170). Upper bounds on the word and symbol error probabilities for soft-decision decoding are again given by (2-165) and (2-166). A comparison of (2-163) and (2-173) indicates that for large values of $\bar{\gamma}_s$ and the same block code, PSK and QPSK have a 3 dB advantage over coherent binary FSK in a fading environment.

As in Section 2.3, the preceding analysis can be extended to Nakagami fading if the parameter m is a positive integer. It is found that the preceding equations for the error probabilities remain valid except that w in (2-162) is replaced by mw and p is given by (2-84) or (2-85) with $\bar{\gamma} = r\bar{\gamma}_b$.

When fast fading makes it impossible to obtain accurate estimates of the $\{\alpha_i\}$ and $\{\theta_i\}$, noncoherent MFSK is a suitable modulation. Expanding the argument of the exponential function in (2-171), assuming that θ_i is uniformly distributed over $[0, 2\pi)$, expressing y_{ki} in polar form, observing that the integral over θ_i is over one period of the integrand, and using the identity (A-30), we obtain the conditional probability density function of y_{ki} given l and α_i:

$$f(y_{ki}|l, \alpha_i) = \frac{1}{4\pi\mathcal{E}_s N_0} \exp\left[-\frac{|y_{ki}|^2 + 4\mathcal{E}_s^2\alpha_i^2\delta_{k\nu}}{4\mathcal{E}_s N_0}\right] I_0\left(\frac{\alpha_i|y_{ki}|\delta_{k\nu}}{N_0}\right) \tag{2-174}$$

Assuming that each α_i is statistically independent and has the same Rayleigh probability density function given by (1-20), $f(y_{ki}|l)$ can be evaluated by using the identity (A-33). Calculating the log-likelihood function and eliminating irrelevant terms and factors, we obtain the *maximum-likelihood metrics for noncoherent MFSK*:

$$U(l) = \sum_{i=1}^{n} R_{li}^2, \quad l = 1, 2, \ldots, q^k \tag{2-175}$$

where $R_{li} = |y_{\nu i}|$ denotes the envelope produced by the filter matched to the transmitted signal for symbol i of codeword l.

Assuming that the all-zero codeword was transmitted, a derivation similar to the one preceding (2-113) again verifies (2-162), (2-165), and (2-166) with

$$p = \frac{1}{2 + \bar{\gamma}_s} \quad \text{(noncoherent MFSK)} \tag{2-176}$$

where $\bar{\gamma}_s$ is given by (2-170). A comparison of (2-163) and (2-176) indicates that for large values of $\bar{\gamma}_b$ and the same block code, PSK and QPSK have

an approximate 6 dB advantage over noncoherent binary FSK in a fading environment. Thus, the fading accentuates the advantage that exists for the AWGN channel.

A comparison of (2-64) with (2-162) and the first term on the right-hand side of (2-166) indicates that a binary block code with maximum-likelihood decoding provides an equivalent diversity equal to d_m if $P_{ib} = P_{is}$ is low enough that the first term in (2-166) dominates and $d_m A_{d_m}/n$ is sufficiently small.

For hard-decision decoding, the symbol error probability is determined by setting $P_s = p$, where p is given by (2-163) for coherent PSK, (2-173) for coherent MFSK, (2-176) for noncoherent MFSK, or (2-101) for DPSK. The word error probability is upper bounded by [3], [6], [10]

$$P_w \leq \sum_{i=t+1}^{n} \binom{n}{i} P_s^i (1-P_s)^{n-i} \tag{2-177}$$

where t is the number of symbol errors that a bounded-distance decoder can correct. For tightly packed codes that have few decoding failures [14], the information-bit error probability is

$$P_{ib} \approx \frac{q}{2(q-1)} \left[\sum_{i=t+1}^{d_m} \frac{d_m}{n} \binom{n}{i} P_s^i (1-P_s)^{n-i} + \sum_{i=d_m+1}^{n} \binom{n-1}{i-1} P_s^i (1-P_s)^{n-i} \right] \tag{2-178}$$

For loosely packed codes, (2-178) provides an approximate upper bound on P_{ib}.

Figure 22 illustrates P_{ib} for an extended Golay (24,12) code with $L = 1$ and MRC with $L = 1, 4, 5$, and 6 diversity branches. A Rayleigh fading channel and binary PSK are assumed. The extended Golay (24,12) code has 12 information bits, $r = 1/2$, $d_m = 8$, and $t = 3$. The numbers of codeword weights in (2-162) are $A_8 = 759$, $A_{12} = 2576$, $A_{16} = 759$, $A_{24} = 1$, and $A_l = 0$, otherwise. The MRC plots assume that a single bit is transmitted. The SNR per code symbol $\bar{\gamma}_s = \bar{\gamma}/2$, where $\bar{\gamma}$ is the average SNR per bit and branch. The figure indicates the benefits of coding particularly when the desired P_{ib} is low. At $P_{ib} = 10^{-3}$, the (24,12) code with hard decisions provides on 11 dB advantage over uncoded PSK; with soft decisions, the advantage becomes 16 dB. The advantage of soft-decision decoding relative to hard-decision decoding increases to more than 10 dB at $P_{ib} = 10^{-7}$, a vast gain over the approximately 2 dB advantage of soft-decision decoding for the AWGN channel. At $P_{ib} = 10^{-9}$, the (24,12) code with soft decisions outperforms MRC with $L = 5$ and is nearing the performance of MRC with $L = 6$. However, since $A_{d_m} = A_8 = 759$, the equivalent diversity will not reach $L = 8$ even for very low P_{ib}. For noncoherent binary FSK, all the curves in the figure are shifted approximately 6 dB to the right when $P_{ib} \leq 10^{-3}$.

Since the soft-decision decoding of long block codes is usually impractical, convolutional codes are more likely to give a good performance over a fading channel. The metrics are basically the same as they are for block codes

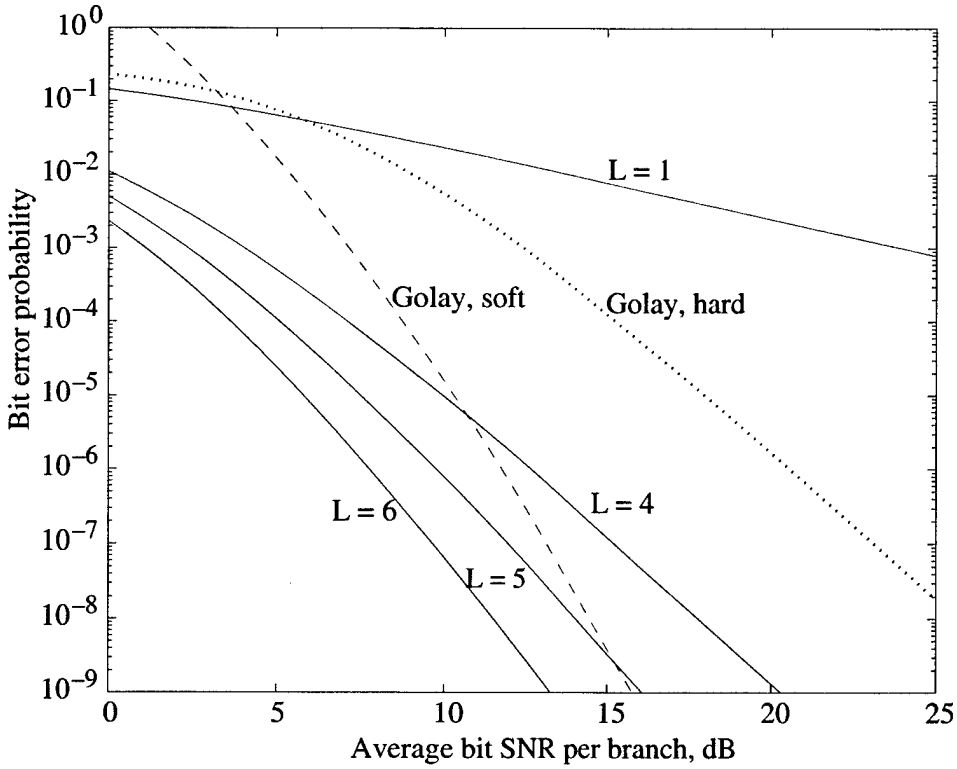


Figure 22. Information-bit error probability for extended Golay (24,12) code with soft and hard decisions, coherent PSK modulation, and Rayleigh fading, and for maximal-ratio combining with $L = 1, 4, 5$, and 6.

with the same modulation, but they are evaluated over path segments that diverge from the correct path through the trellis and then merge with it subsequently. The linearity of binary convolutional codes ensures that all-zero path can be assumed to be the correct one when calculating the decoding error probability. Let l denote the Hamming distance of an incorrect path from the correct all-zero path. If perfect symbol interleaving is used, then the probability of error in the pairwise comparison of two paths with an unmerged segment is $P_2(l)$, which is given by (2-162). The probability of an information-bit error in soft-decision decoding is upper bounded by [10]

$$P_{ib} \leq \frac{1}{k} \sum_{l=d_f}^{\infty} B(l) P_2(l) \tag{2-179}$$

where $B(l)$ is the number of information-bit errors over all paths with unmerged segments at Hamming distance l, k is the number of information bits per trellis branch, and d_f is the minimum free distance, which is the minimum Hamming distance between any two convolutional codewords. This upper bound approaches $B_{d_f} P_2(d_f)/k$ as $P_{ib} \to 0$ so the equivalent diversity is d_f if P_{ib} and $B(d_f)/k$ are small.

In general, d_f increases with the constraint length of the convolutional code. However, if each encoder output bit is repeated n_r times, then the minimum distance of the convolutional code increases to $n_r d_f$ without a change in the constraint length, but at the cost of a bandwidth expansion by the factor η_r. The information-bit error probability has the upper bound given by

$$P_{ib} \leq \frac{1}{k} \sum_{l=d_f}^{\infty} B(l) P_2(n_r l) \tag{2-180}$$

Figure 23 illustrates P_{ib} as a function of $\bar{\gamma}_b$ for the Rayleigh-fading channel and binary convolutional codes with different values of the constraint length K, the code rate r, and the number of repetitions n_r. Equations (2-180) and (2-162) with $k = 1$ are used, and the $\{B(l)\}$ are taken from the listings for seven terms in [15]. The figure indicates that an increase in the constraint length provides a much greater performance improvement for the Rayleigh-fading channel than the increase does for the AWGN channel [16]. For a fixed constraint length, the rate-1/4 codes give a better performance than the rate-1/2 codes with $n_r = 2$, which require the same bandwidth but are less complex to implement. The latter two codes require twice the bandwidth of the rate-1/2 code with no repetitions.

The issues are similar for trellis-coded modulation [3], [6], [10], which provides a coding gain without a bandwidth expansion. However, if parallel state transitions occur in the trellis, then $d_f = 1$, which implies that the code provides no diversity protection against fading. Thus, for fading communications, a conventional trellis code with distinct transitions from each state to all other states must be selected. Since Rayleigh fading causes large amplitude variations, multiphase PSK is usually a better choice than multilevel quadrature amplitude modulation (QAM) for the symbol modulation.

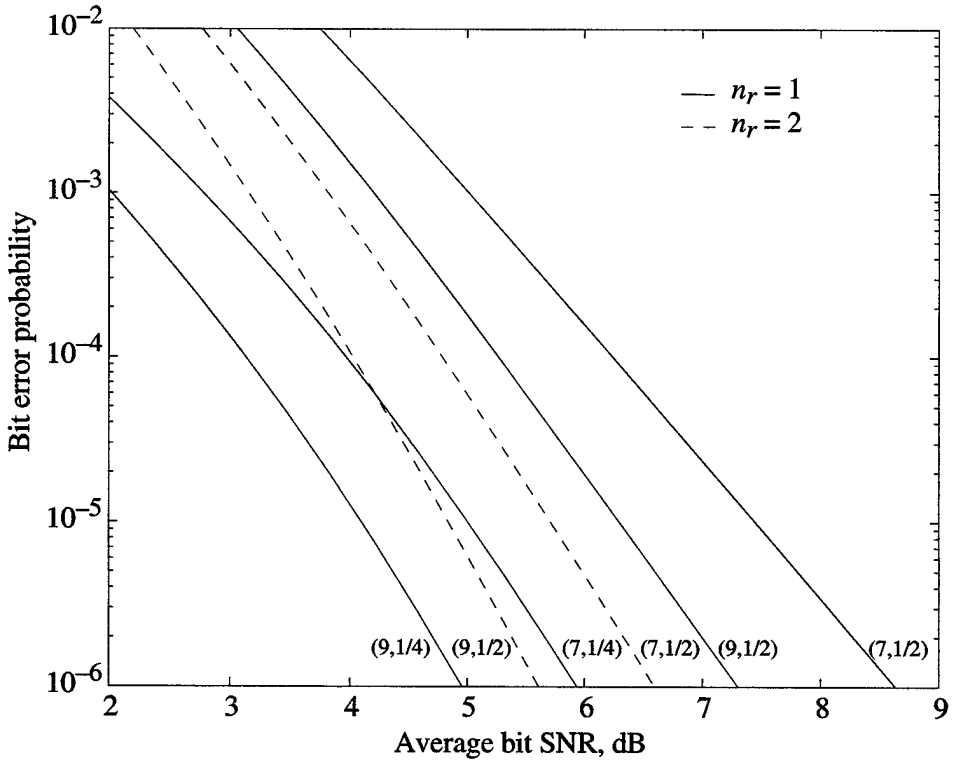


Figure 23. Information-bit error probability for Rayleigh fading, coherent PSK, and binary convolutional codes with various values of (K, r) and n_r.

However, the optimum trellis decoder uses coherent detection and requires an estimate of the channel attenuation.

Whether a block, convolutional, or trellis code is used, the results of this section indicate that the minimum Hamming distance rather than the minimum Euclidean distance is the critical parameter in designing an effective code for the Rayleigh fading channel.

Turbo codes or serial concatenated codes with iterative decoding based on the *maximum a posteriori* criterion can provide excellent performance. However, the system must be able to accommodate considerable decoding delay and computational complexity. Even without iterative decoding, a serial concatenated code with an outer Reed-Solomon code and an inner binary convolutional code [3], [10] can be effective against Rayleigh fading. In the worst case, each output bit error of the inner decoder causes a separate symbol error at the input to the Reed-Solomon decoder. Therefore, the input symbol error probability P_{s1} is upper bounded by m times the bit error probability at the inner-decoder output, where $m = \log_2 q$ is the number of bits in a Reed-Solomon code symbol. Inequality (2-179) implies that

$$P_{s1} \leq \frac{m}{k} \sum_{l=d_f}^{\infty} B(l) P_2(l) \tag{2-181}$$

Assuming that symbol deinterleaving ensures independent symbol errors at the Reed-Solomon decoder input, an upper bound on P_{ib} is determined from (2-178) with P_{s1} substituted in place of P_s. The concatenated code has a code rate

$$r = r_1 r_0 \tag{2-182}$$

where r_1 is the inner-code rate and r_0 is the outer-code rate. For coherent PSK modulation with soft-decision decoding, $P_2(l)$ is given by (2-162), p is given by (2-163), and $\bar{\gamma}_s$ is given by (2-164) and (2-182).

Figure 24 depicts examples of the upper bound on P_{ib} as a function $\bar{\gamma}_b$ for Rayleigh fading, coherent PSK, soft decisions, an inner binary convolutional code with $K = 7$, $r_1 = 1/2$, and $k = 1$, and various Reed-Solomon (n, k) outer codes. The required bandwidth is B_u/r, where B_u is the uncoded PSK bandwidth. Thus, the codes of the figure require a bandwidth less than $3B_u$.

2.8 Space-Time Coding

Multiple antennas at a transmitter can be used to provide *transmit diversity*, which is analogous to the spatial diversity provided by multiple antennas at a receiver, but there are two principal differences. One difference is that arriving signals must somehow be separated at each receiving antenna and then combined appropriately. Another difference is that the total transmission power must be divided among the diversity signals. Signal separation

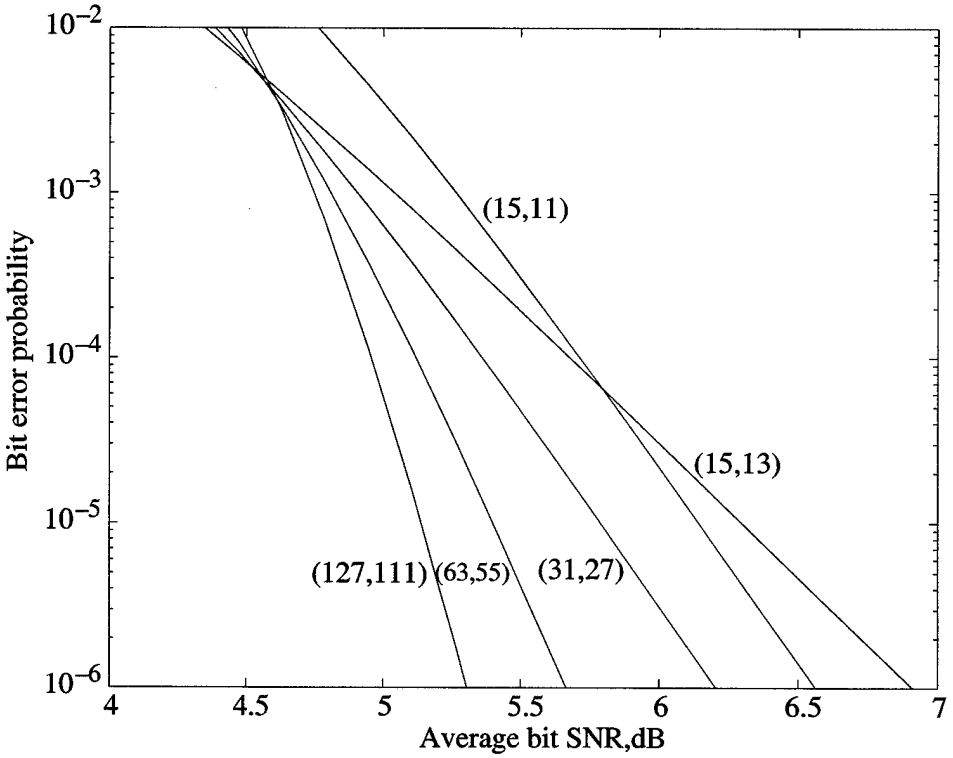


Figure 24. Information-bit error probability for Rayleigh fading, coherent PSK, soft decisions, and concatenated codes comprising an inner binary convolutional code with $K = 7$ and $r_1 = 1/2$, and various Reed-Solomon (n, k) outer codes.

may be facilitated by using multiple delayed versions of a symbol, feedforward or training information to enable channel estimation, selection diversity with feedback from the receiver to the transmitter, orthogonal channel coding of the signals, or other methods.

A much more powerful approach to improving system performance is to use both multiple transmit and multiple receive antennas in a *multiple-input, multiple-output* configuration. A dramatic increase in the data rate or system capacity is possible by using *space-time codes*, which are codes that introduce both temporal and spatial diversity into signals transmitted from different antennas [17]. Space-time codes with multiple transmit antennas, which have been specified in third-generation cellular standards, provide a favorable tradeoff between power consumption and spectral efficiency.

3. Direct-Sequence Code-Division Multiple Access

Multiple access is the ability of many users to communicate with each other while sharing a common transmission medium. Wireless multiple-access communications are possible if the transmitted signals are orthogonal or separable in some sense. Signals may be separated in time (*time-division multiple access* or TDMA), frequency (*frequency-division multiple access* or FDMA), or code (*code-division multiple access* or CDMA). CDMA is realized by using spread-spectrum modulation while transmitting signals from multiple users in the same frequency band at the same time. All signals use the entire allocated spectrum. Information theory indicates that in an isolated cell, CDMA systems achieve the same spectral efficiency as TDMA or FDMA systems only if optimal multiuser detection is used. However, even with single-user detection, CDMA is advantageous for cellular networks because it eliminates the need for frequency and timeslot coordination among cells and allows carrier-frequency reuse in adjacent cells. Frequency planning is vastly simplified. A major CDMA advantage exists in networks accommodating voice communications. A voice-activity detector activates the transmitter only when the user is talking. Since typically fewer than 40% of the users are talking at any given time, the number of telephone users can be increased while maintaining a specified average interference power. Another major CDMA advantage is the ease with which it can be combined with multibeamed antenna arrays that are either adaptive or have fixed patterns covering cell sectors. There is no practical means of reassigning time slots in TDMA systems or frequencies in FDMA systems to increase capacity by exploiting intermittent voice signals or multibeamed arrays. Reassignments to accommodate variable data rate are almost always impractical in FDMA or TDMA systems. These general advantages of CDMA, combined with the resistance of spread-spectrum signals to jamming, interception, and multipath interference, make CDMA the most attractive choice for most mobile communications

The two principal types of spread-spectrum CDMA are *direct-sequence* CDMA and *frequency-hopping* CDMA. In direct-sequence CDMA, pulses with a large bandwidth relative to the symbol rate can be generated by using a chip waveform that is modulated by a spreading sequence. In frequency-hopping CDMA, the carrier frequency of a transmission is changed periodically.

3.1 Direct-Sequence System with Coherent PSK

A received direct-sequence signal with PSK data modulation can be represented by

$$s(t) = \sqrt{2S}m(t)p(t)\cos 2\pi f_0 t \tag{3-1}$$

where S is the average power, $m(t)$ is the data modulation, $p(t)$ is the spreading waveform, and f_0 is the carrier frequency. The data modulation is a sequence of nonoverlapping rectangular pulses, each of which has an amplitude equal to +1 or –1. Each pulse of $m(t)$ represents a data symbol and has a duration of T_s. The spreading waveform has the form

$$p(t) = \sum_{i=-\infty}^{\infty} p_i \psi(t - iT_c) \tag{3-2}$$

where p_i is equal to +1 or –1 and represents one chip of a pseudonoise sequence $\{p_i\}$. The chip waveform $\psi(t)$ has duration T_c that is shorter than T_s. It is convenient, and entails no loss of generality, to normalize the energy content of the chip waveform according to

$$\frac{1}{T_c} \int_0^{T_c} \psi^2(t)dt = 1 \tag{3-3}$$

Because the transitions of a data symbol and the chips coincide on both sides of a symbol, the *processing gain*, defined as

$$G = \frac{T_s}{T_c} \tag{3-4}$$

is an integer equal to the number of chips in a symbol interval.

In the presence of white Gaussian noise, the optimum demodulator for the detection of a single symbol is the correlator shown in Figure 25. This correlator is optimal only if the interference is white Gaussian noise, but is a reasonable approach against other types of interference. An equivalent matched-filter demodulator combines the sampled outputs of a chip matched filter with a stored pseudonoise sequence. However, the matched-filter implementation is not practical for a long sequence that extends over many data symbols

In the subsequent analysis, perfect phase, sequence, and symbol synchronization are assumed. The effects of the initial wideband filtering are assumed to be negligible. The total received signal is

$$r(t) = s(t) + i(t) + n(t) \tag{3-5}$$

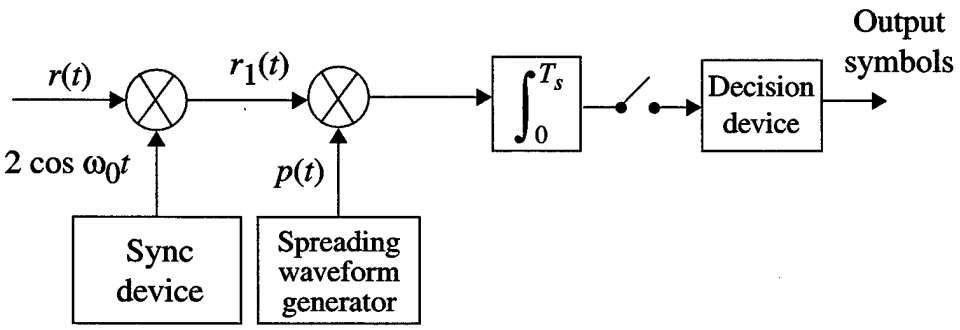


Figure 25. Basic elements of receiver for direct-sequence signal with coherent PSK.

where $i(t)$ is the interference, and $n(t)$ denotes the zero-mean white Gaussian noise. Assuming that synchronization has been established, the input sample applied to the decision device at the end of a data-symbol interval defined as $[0,T_s]$ is

$$V = \int_0^{T_s} 2r(t)p(t)\cos 2\pi f_0 t \, dt \tag{3-6}$$

In the evaluation of this equation, it is assumed that $f_0 >> 1/T_s$ so that the integral over a double-frequency term is negligible. The preceding equations and the fact that $m(t)$ is constant over the integration interval yield the *decision variable*

$$\begin{aligned} V &= \sqrt{2S}\int_0^{T_s} m(t)\sum_{i=0}^{G-1}\psi^2(t-iT_c)dt + V_1 + V_2 \\ &= \pm\sqrt{2S}T_s + V_1 + V_2 \end{aligned} \tag{3-7}$$

where

$$V_1 = \int_0^{T_s} 2i(t)p(t)\cos 2\pi f_0 t \, dt \tag{3-8}$$

$$V_2 = \int_0^{T_s} 2n(t)p(t)\cos 2\pi f_0 t \, dt \tag{3-9}$$

The factor $p(t)$ in $i(t)p(t)$ ensures that the interference energy is spread over a wide band.

Suppose that $m(t)$ = +1 represents the logic symbol 1 and $m(t) = -1$ represents the logic symbol 0. The decision device produces the symbol 1 if $V > 0$ and the symbol 0 if $V < 0$. An error occurs if $V < 0$ when $m(t)$ = +1 or if $V > 0$ when $m(t)$ = –1. The probability that $V = 0$ is zero.

The white Gaussian noise has autocorrelation

$$R_n(\tau) = \frac{N_0}{2}\delta(t-\tau) \tag{3-10}$$

where $N_0/2$ is the two-sided noise power spectral density. A straightforward calculation using (3-9) and (3-10) and assuming that $f_0 >> 1/T_s$ yields

$$\text{var}(V_2) = N_0 T_s \tag{3-11}$$

Suppose that at a particular receiver, a multiple-access interference signal has the form

$$i(t) = \sqrt{2I}m_q(t)q(t)\cos(2\pi f_0 t + \phi) \tag{3-12}$$

where $m_q(t)$ is the data modulation, $q(t)$ is the spreading waveform, I is the received power, and f_0 is the same for all systems in a CDMA network.

Equation (3-8) implies that the resulting interference term at the input of the decision device has the form

$$V_1 = \sqrt{2I}\cos\phi \int_0^{T_s} m_q(t)q(t)p(t)\,dt \tag{3-13}$$

where a double-frequency term is neglected and $p(t)$ is the spreading waveform of the desired signal. Suppose that the communication signals are *synchronous* so that data-symbol transitions of $p(t)$ and $q(t)$ are aligned at the receiver. Then $m_q(t)$ is constant over $[0,T_s]$. If all the spreading sequences have a common period equal to T_s, the common data-symbol duration, then V_1 is proportional to the cross-correlation between $q(t)$ and $p(t)$, which is defined as

$$C_{pq} = \frac{1}{T_s}\int_0^{T_s} q(t)p(t)dt \tag{3-14}$$

If both $p(t)$ and $q(t)$ have the same form given by (3-2) and have the same chip waveform, then (3-14), (3-3), and (3-4) yield

$$C_{pq} = \frac{1}{G}\sum_{i=1}^{G} p_i q_i = \frac{A-D}{G} \tag{3-15}$$

where both p_i and q_i are equal to either +1 or -1, A denotes the number of terms with $p_i q_i = +1$, and D denotes the number of terms with $p_i q_i = -1$. Each level of p_i or q_i represents a binary digit a_i or b_i, respectively, where $a_i = (p_i + 1)/2$ and $b_i = (q_i + 1)/2$. Therefore, A equals the number of agreements in the corresponding digits of the sequences $\{a_i\}$ and $\{b_i\}$, and D equals the number of disagreements in the corresponding digits. The sequences are *orthogonal* if $C_{pq} = 0$. If the sequences are orthogonal, then $V_1 = 0$ and the multiple-access interference $i(t)$ is suppressed at the receiver. A large number of multiple-access interference signals can be suppressed in a network if each such signal has its data-symbol transitions aligned and the spreading sequences are mutually orthogonal.

Two sequences of digits, each of length two, are orthogonal if each sequence is described by one of the rows of the 2×2 matrix

$$\mathbf{H}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{3-16}$$

because $A = D = 1$. A set of 2^n sequences, each of length 2^n, is obtained by using the rows of the matrix

$$\mathbf{H}_n = \begin{bmatrix} \mathbf{H}_{n-1} & \mathbf{H}_{n-1} \\ \mathbf{H}_{n-1} & \bar{\mathbf{H}}_{n-1} \end{bmatrix}, \quad n = 2, 3, \ldots \tag{3-17}$$

where $\bar{\mathbf{H}}_{n-1}$ is the *complement* of $\mathbf{H}_{n-1}$, obtained by replacing each 1 and 0 by 0 and 1, respectively, and $\mathbf{H}_1$ is defined by (3-16). Any pair of rows in $\mathbf{H}_n$ differ in exactly 2^{n-1} columns, thereby ensuring orthogonality of the corresponding sequences. The $2^n \times 2^n$ matrix $\mathbf{H}_n$, which is called a *Hadamard* matrix, can be used to generate all the spreading sequences needed for synchronous direct-sequence communications.

3.2 Sequences with Small Cross-Correlations

When the symbol transitions of *asynchronous* multiple-access signals at a receiver are not simultaneous, usually because of changing path-length differences among the various communication links, the receiver performance depends on the cross-correlations of spreading sequences that are shifted relative to each other. Sets of periodic sequences with small cross-correlations are desirable to limit the effect of multiple-access interference. Maximal-length sequences [10], [18] which have the longest periods of sequences generated by a linear feedback shift register of fixed length, are often inadequate.

One set of periodic sequences with small cross-correlations consists of the *Gold sequences*, which may be generated by the modulo-2 addition of certain pairs of maximal-length sequences called *preferred pairs* [19]. One form of a Gold sequence generator is shown in Figure 26. If each maximal-length sequence generator has m stages, different Gold sequences in a set or code are generated by selecting the initial state of one maximal-length sequence generator and then shifting the initial state of the other generator. Since any shift from 0 to $2^m - 2$ results in a different Gold sequence, $2^m - 1$ different Gold sequences can be produced by the system of Figure 26. Gold sequences identical to maximal-length sequences are produced by setting the state of one of the maximal-length sequence generators to zero. Altogether, there are $2^m + 1$ different Gold sequences, each with a period of $2^m - 1$, in the set.

The *periodic cross-correlation* function between periodic sequences $\{p_i\}$ and $\{q_i\}$ with the same length G is defined as

$$\theta_{pq}(k) = \frac{1}{G}\sum_{i=1}^{G} p_i q_{i+k}\ ,\ k = 1, 2, \ldots, G-1 \qquad (3\text{-}18)$$

The periodic cross-correlation function between any two Gold sequences in a set with period $G = 2^m - 1$ can only take the values $-t(m)/G, -1/G$, or

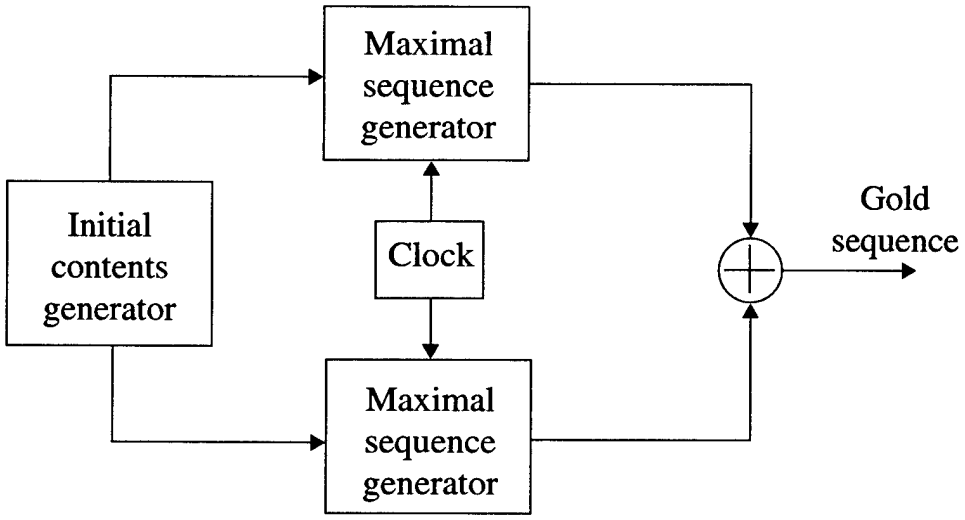


Figure 26. Gold sequence generator.

$[t(m)-2]/G$, where

$$t(m) = \begin{cases} 2^{(m+1)/2} + 1 \ , & m \text{ is odd} \\ 2^{(m+2)/2} + 1 \ , & m \text{ is even} \end{cases} \tag{3-19}$$

Aside from its peak value of unity, an autocorrelation function also takes only these values. The peak magnitude of the periodic cross-correlation function between any two Gold sequences is

$$\theta_p = \frac{t(m)}{G} \tag{3-20}$$

which is often much smaller than what is possible with maximal sequences.

A set of $2^{m/2}$ *Kasami sequences* with period $2^m - 1$ can be generated from a maximal-length sequence if m is even [19]. The periodic cross-correlation function between any two Kasami sequences with period $G = 2^m - 1$ can only take the values $-s(m)/G$, $-1/G$, or $[s(m)-2]/G$, where

$$s(m) = 2^{m/2} + 1 \tag{3-21}$$

The peak magnitude of the periodic cross-correlation function between any two Kasami sequences is

$$\theta_p = \frac{s(m)}{G} = \frac{2^{m/2} + 1}{2^m - 1} \tag{3-22}$$

The Kasami sequences are optimal in the sense that θ_p has the minimum value for any set of sequences of the same size and period.

As an example, let m = 11. It can be shown [18] that there are 176 maximal sequences of period $2^{11} - 1 = 2047$. The peak cross-correlation for the set is $\theta_p = 0.14$. A much larger set of $2^{11} + 1 = 2049$ Gold sequences of period 2047 has peak cross-correlation given by (3-19), which yields $\theta_p = 0.03$. If m = 10, there are 60 maximal sequences, 1025 Gold sequences, and 32 Kasami sequences with period 1023. The peak cross-correlations are 0.37, 0.06, and 0.03, respectively.

If all the spreading sequences in a network of asynchronous CDMA systems have a common period equal to data-symbol duration, then by the proper selection of the sequences and their relative phases, one can obtain a system performance better than that theoretically attainable with random sequences. However, the number of suitable sequences is too small for many applications, and long sequences that extend over many channel symbols provide more system security.

3.3 Direct-Sequence Systems with PSK and Random Sequences

It is natural and analytically desirable to model a long spreading sequence as a random binary sequence. The random-binary-sequence model does not seem to obscure important exploitable characteristics of long sequences

and is a reasonable approximation even for short sequences in networks with asynchronous communications. A random binary sequence consists of statistically independent symbols, each of which takes the value +1 with probability $\frac{1}{2}$ or the value -1 with probability $\frac{1}{2}$. Thus, $E[p_i] = E[p(t)] = 0$. It then follows from (3-7) to (3-9) that $E[V_1] = E[V_2] = 0$, and the mean value of the decision variable is

$$E[V] = \pm\sqrt{2S}\, T_s \tag{3-23}$$

in the direct-sequence system with coherent PSK. Substituting (3-2) into (3-8), we obtain

$$V_1 = \sum_{v=0}^{G-1} p_\nu J_\nu \tag{3-24}$$

where

$$J_\nu = \int_{\nu T_c}^{(\nu+1)/T_c} 2i(t)\psi(t - \nu T_c)\cos 2\pi f_0 t\, dt \tag{3-25}$$

Consider a network of direct-sequence systems, each of which uses a random spreading sequence. A suitable model for multiple-access interference is given by

$$i(t) = \sum_{i=1}^{K} \sqrt{2I_i}\, q_i(t - \tau_i)\cos(2\pi f_0 t + \phi_i) \tag{3-26}$$

where I_i is the average power at the receiver due to interference signal i, K is the number of interfering spread-spectrum signals, $q_i(t)$ is the spreading waveform of signal i, τ_i is the relative delay of signal i, and ϕ_i is the phase angle of signal i including the effect of carrier time delay. Each spreading waveform has the form

$$q_i(t) = \sum_{j=-\infty}^{\infty} q_{ij}\psi(t - jT_c) \tag{3-27}$$

where the chip waveforms are assumed to be identical throughout the network and each $\{q_{ij}\}$ is a sequence modeled as a random binary sequence that is independent of $\{p_v\}$. Since the data modulation in an interference signal is modeled as a random binary sequence, it can be subsumed into $\{q_{ij}\}$ with no loss of generality. The model for $q_i(t)$ implies that only time delays modulo-T_c are significant and, thus, we can assume that $0 \le \tau_i < T_c$ without loss of generality.

Substituting (3-26) and (3-27) into (3-25), using the time-limited character of $\psi(t)$, and neglecting the double-frequency term, we obtain

$$\begin{aligned} J_\nu = \sum_{i=1}^{K} \sqrt{2I_i}\cos\phi_i \Bigg\{ & q_{i,\nu-1}\int_{\nu T_c}^{\nu T_c+\tau_i} \psi(t - \nu T_c)\psi[t - (\nu-1)T_c - \tau_i]dt \\ & + q_{i,\nu}\int_{\nu T_c+\tau_i}^{(\nu+1)T_c} \psi(t - \nu T_c)\psi(t - \nu T_c - \tau_i)dt \Bigg\} \end{aligned} \tag{3-28}$$

We define the *partial autocorrelation* for the chip waveform as

$$R_\psi(s) = \int_0^s \psi(t)\psi(t + T_c - s)dt, \quad 0 \le s < T_c \tag{3-29}$$

Substitution into (3-28) and appropriate changes of variables in the integrals yield

$$J_\nu = \sum_{i=1}^{K} \sqrt{2I_i} \cos\phi_i [q_{i,\nu-1} R_\psi(\tau_i) + q_{i\nu} R_\psi(T_c - \tau_i)] \tag{3-30}$$

It is assumed in this section that K and the $\{I_i\}$ are constants.

For rectangular pulses in the spreading waveform,

$$\psi(t) = \begin{cases} 1 \ , & 0 \le t \le T_c \\ 0 \ , & \text{otherwise.} \end{cases} \tag{3-31}$$

Consequently,

$$R_\psi(s) = s, \quad \text{rectangular pulse} \tag{3-32}$$

For sinusoidal pulses in the spreading waveform,

$$\psi(t) = \begin{cases} \sqrt{2}\sin\left(\dfrac{\pi}{T_c}t\right) \ , & 0 \le t \le T_c \\ 0 \qquad\quad , & \text{otherwise.} \end{cases} \tag{3-33}$$

Substituting this equation into (3-29), using a trigonometric identity, and performing routine integrations, we obtain

$$R_\psi(s) = \frac{T_c}{\pi}\sin\left(\frac{\pi}{T_c}s\right) - s\cos\left(\frac{\pi}{T_c}s\right), \quad \text{sinusoidal pulse} \tag{3-34}$$

Since both J_ν and $J_{\nu+1}$ contain the same random variable, $q_{i\nu}$, it does not appear at first that the terms in (3-24) are statistically independent even when $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_k)$ and $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_k)$ are given. Thus, the following lemma is needed [20].

Lemma. Suppose that $\{\alpha_i\}$ and $\{\beta_i\}$ are statistically independent, random binary sequences. Let x and y denote arbitrary constants. Then $\alpha_i\beta_j x$ and $\alpha_i\beta_k y$ are statistically independent random variables when $j \ne k$.

Proof. Let $P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b)$ denote the joint probability that $\alpha_i\beta_j x = a$ and $\alpha_i\beta_k y = b$ where $|a| = |x|$ and $|b| = |y|$. From the theorem of total probability, it follows that

$$\begin{aligned} &P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b) \\ &= P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b, \alpha_i = 1) + P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b, \alpha_i = -1) \\ &= P(\beta_j x = a, \beta_k y = b, \alpha_i = 1) + P(\beta_j x = -a, \beta_k y = -b, \alpha_i = -1) \end{aligned}$$

From the independence of $\{\alpha_i\}$ and $\{\beta_j\}$ and the fact that they are random binary sequences, we obtain the further simplification for $j \neq k$, $x \neq 0$, and $y \neq 0$:

$$\begin{aligned}
&P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b) \\
&= P(\beta_j x = a)P(\beta_k y = b)P(\alpha_i = 1) + P(\beta_j x = -a)P(\beta_k y = -b)P(\alpha_i = -1) \\
&= \frac{1}{2}P\left(\beta_j = \frac{a}{x}\right)P\left(\beta_k = \frac{b}{y}\right) + \frac{1}{2}P\left(\beta_j = -\frac{a}{x}\right)P\left(\beta_k = -\frac{b}{y}\right)
\end{aligned}$$

Since β_j equals +1 or -1 with equal probability, $P(\beta_j = a/x) = P(\beta_j = -a/x)$ and thus

$$\begin{aligned}
P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b) &= P\left(\beta_j = \frac{a}{x}\right)P\left(\beta_k = \frac{b}{y}\right) \\
&= P(\beta_j x = a)P(\beta_k y = b)
\end{aligned}$$

A similar calculation gives

$$P(\alpha_i\beta_j x = a)P(\alpha_i\beta_k y = b) = P(\beta_j x = a)P(\beta_k y = b)$$

Therefore,

$$P(\alpha_i\beta_j x = a, \alpha_i\beta_k y = b) = P(\alpha_i\beta_j x = a)P(\alpha_i\beta_k y = b)$$

which satisfies the definition of statistical independence of $\alpha_i\beta_j x$ and $\alpha_i\beta_k y$. The same relation is trivial to establish for $x = 0$ or $y = 0$. □

A direct application of the lemma indicates that when ϕ and τ are given, the terms in (3-24) are statistically independent. Since $p_\nu^2 = 1$, the conditional variance is

$$\text{var}(V_1) = \sum_{\nu=0}^{G-1} \text{var}(J_\nu) \tag{3-35}$$

The independence of the K spreading sequences, the independence of successive terms in each random binary sequence, and (3-30) imply that the conditional variance of J_ν is independent of ν and, therefore,

$$\text{var}(V_1) = \sum_{i=1}^{K} 2GI_i \cos^2\phi_i[R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i)] \tag{3-36}$$

Since the terms of V_1 in (3-24) are independent, zero-mean random variables that are uniformly bounded and $\text{var}(V_1) \to \infty$ as $G \to \infty$, the central limit theorem [21] implies that $V_1/\sqrt{\text{var}(V_1)}$ converges in distribution to a Gaussian random variable with mean 0 and variance 1. Thus, when ϕ and τ are given, the conditional distribution of V_1 is approximately Gaussian when G is large. Since V_2 has a Gaussian distribution and is independent of V_1, (3-7) implies that V has an approximate Gaussian distribution with mean given by (3-23), and $\text{var}(V) = \text{var}(V_1) + \text{var}(V_2)$.

Under the Gaussian approximation, a straightforward derivation using the Gaussian distribution of the decision statistic V indicates that the conditional symbol error probability given ϕ and τ is

$$P_s(\phi,\tau) = Q\left(\sqrt{\frac{2E_s}{N_{0e}(\phi,\tau)}}\right) \tag{3-37}$$

where $E_s = ST_s$ is the *energy per symbol* in $m(t)$, and the *equivalent noise power spectral density* is defined as

$$N_{0e}(\phi,\tau) = N_0 + \sum_{i=1}^{K} 2\frac{I_i}{T_c}\cos^2\phi_i[R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i)] \tag{3-38}$$

For rectangular pulses, this equation simplifies to

$$N_{0e}(\phi,\tau) = N_0 + \sum_{i=1}^{K} 2\frac{I_i}{G}T_s\cos^2\phi_i\left(1 - 2\frac{\tau_i}{T_c} + 2\frac{\tau_i^2}{T_c^2}\right) \tag{3-39}$$

which shows explicitly how the interference power I_i is reduced by the processing gain G. Numerical evaluations [20] give strong evidence that the error in (3-37) due to the Gaussian approximation is negligible if $G \geq 50$.

To determine the symbol error probability at the output of the decision device, we calculate the expected value of $P_s(\phi,\tau)$ over the distributions of ϕ and τ. For an asynchronous network, it is assumed that the time delays are independent and uniformly distributed over [0,T_c) and that the phase angles $\theta_i, i = 1, 2, \ldots, K$, are uniformly distributed over [0,2π). Therefore, the symbol error probability is

$$P_s = \left(\frac{2}{\pi T_c}\right)^K \int_0^{\pi/2}\cdots\int_0^{\pi/2}\int_0^{T_c}\cdots\int_0^{T_c} P_s(\phi,\tau)d\phi\, d\tau \tag{3-40}$$

where the fact that $\cos^2\phi_i$ takes all its possible values over [0,$\pi/2$) has been used to shorten the integration intervals. Because of the absence of sequence parameters, the amount of computation required for (3-40) is much less than the amount required to compute P_s when the spreading sequence is short. However, the computational requirements of (3-40) are large enough that it is highly desirable to find an accurate approximation that entails less computation.

The conditional symbol error probability given ϕ is defined as

$$P_s(\phi) = \left(\frac{1}{T_c}\right)^K \int_0^{T_c}\cdots\int_0^{T_c} P_s(\phi,\tau)d\tau \tag{3-41}$$

A closed-form approximation to $P_s(\phi)$ greatly simplifies the computation of P_s, which reduces to

$$P_s = \left(\frac{2}{\pi}\right)^K \int_0^{\pi/2}\cdots\int_0^{\pi/2} P_s(\phi)d\phi \tag{3-42}$$

To approximate $P_s(\phi)$, we first obtain upper and lower bounds on it.

For either rectangular or sinusoidal pulses, elementary calculus establishes that

$$R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i) \leq T_c^2 \tag{3-43}$$

Using (3-43) successively in (3-38), (3-37), and (3-41), and performing the trivial integrations that result, we obtain

$$P_s(\phi) \leq Q\left(\sqrt{\frac{E_s}{N_{0u}(\phi)}}\right) \tag{3-44}$$

where

$$N_{0u}(\phi) = N_0 + \sum_{i=1}^{K} 2\frac{I_i}{G} T_s \cos^2 \phi_i \tag{3-45}$$

If $g(\,)$ is a convex function over an interval containing the range of a random variable X, then Jensen's inequality (Appendix C) states that

$$g(E[X]) \leq E[g(X)] \tag{3-46}$$

provided that $E[X]$ exists. To apply (3-46), the successive integrals in (3-41) are interpreted as the evaluation of expected values. For the best results, we set

$$X = R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i). \tag{3-47}$$

Since τ_i is uniformly distributed over $[0, T_c)$, straightforward calculations give

$$E[X] = \frac{1}{T_c}\int_0^{T_c} [R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i)]d\tau_i = hT_c^2 \tag{3-48}$$

where

$$h = \begin{cases} \dfrac{2}{3}, & \text{rectangular pulse} \\ \dfrac{1}{3} + \dfrac{5}{2\pi^2}, & \text{sinusoidal pulse.} \end{cases} \tag{3-49}$$

The function (3-37) has the form

$$g(y) = Q\left(\sqrt{\frac{1}{a + by}}\right) \tag{3-50}$$

where $a > 0$ and $b \geq 0$. Calculating the second derivative of $g(y)$ with respect to y, we find that $g(y)$ is a convex function if $0 \leq a + by \leq \frac{1}{3}$. From

(3-38), (3-43), and the fact that $\cos^2 \phi_i \leq 1$, we obtain a sufficient condition for convexity:

$$E_s \geq \frac{3}{2}\left[N_0 + \sum_{i=1}^{K} 2\frac{I_i}{G}T_s\right] \tag{3-51}$$

Application of Jensen's inequality successively to each component of τ in (3-41) yields

$$P_s(\phi) \geq Q\left(\sqrt{\frac{2E_s}{N_{0l}(\phi)}}\right) \tag{3-52}$$

where

$$N_{0l}(\phi) = N_0 + \sum_{i=1}^{K} 2h\frac{I_i}{G}T_s \cos^2 \phi_i \tag{3-53}$$

If N_0 is negligible, then (3-53) and (3-45) give $N_{0l}/N_{0u} = h$. Thus, a good approximation is provided by

$$P_s(\phi) \approx Q\left(\sqrt{\frac{2E_s}{N_{0a}\phi)}}\right) \tag{3-54}$$

where

$$N_{0a}(\phi) = N_0 + \sum_{i=1}^{K} 2\sqrt{h}\frac{I_i}{G}T_s \cos^2 \phi_i \tag{3-55}$$

If N_0 is negligible, then $N_{0u}/N_{0a} = N_{0a}/N_{0l} = 1/\sqrt{h}$. Therefore, in terms of the value of E_s needed to ensure a given $P_s(\phi)$, the error in using approximation (3-54) instead of (3-41) is bounded by $10\log_{10}(1/\sqrt{h})$ in decibels, which equals 0.88 dB for rectangular pulses and 1.16 dB for sinusoidal pulses. In practice, the error is expected to be only a few tenths of a decibel because $N_0 \neq 0$ and P_s coincides with neither the upper nor the lower bound.

As an example, suppose that rectangular pulses are used, E_s/N_0 = 15 dB, and $K = 1$. Figure 27 illustrates four different evaluations of P_s as a function of $GE_s/IT_s = GS/I$, the *despread signal-to-interference ratio*, which is the signal-to-interference ratio after taking into account the beneficial results from the despreading in the receiver. The accurate approximation is computed from (3-37) and (3-40), the upper bound from (3-44) and (3-42), the lower bound from (3-52) and (3-42), and the simple approximation from (3-54) and (3-42). The figure shows that the accurate approximation moves from the lower bound toward the simple approximation as the symbol error probability decreases. For $P_s = 10^{-5}$, the simple approximation is less than 0.3 dB in error relative to the accurate approximation.

Figure 28 compares the symbol error probabilities for $K = 1$ and $K = 2$, rectangular pulses and E_s/N_0 = 15 dB. The simple approximation is used

for P_s, and the abscissa shows GS/I where I is the interference power of each interfering signal. The figure shows that P_s increases with K, but the shift in P_s is less than 3 dB because the interference signals tend to partially cancel each other.

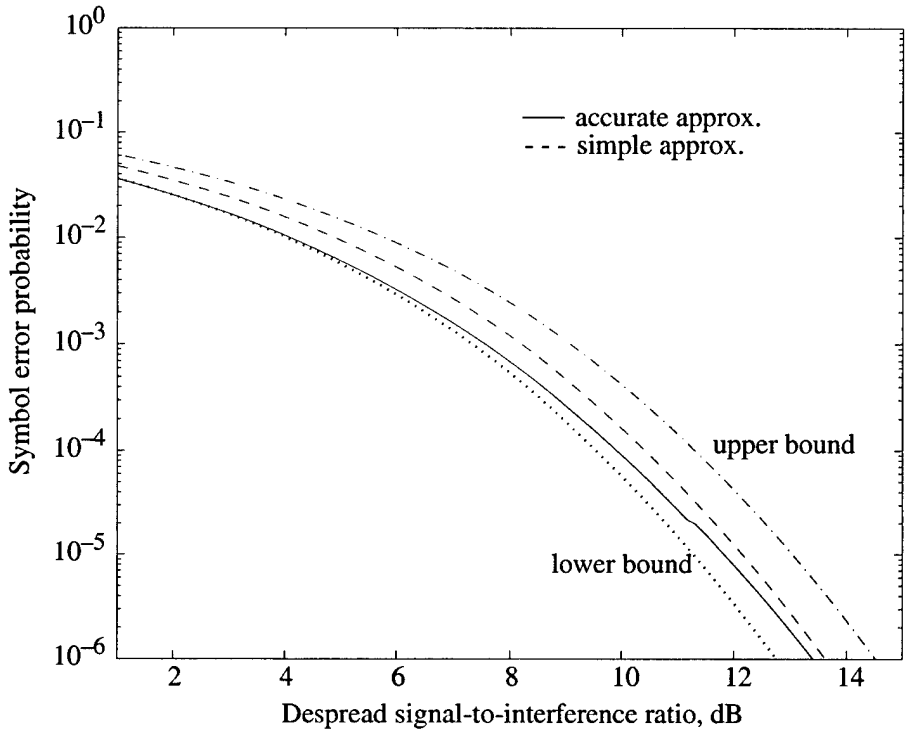


Figure 27. Symbol error probability of direct-sequence system with PSK in presence of single multiple-access interference signal and $E_s/N_0 = 15$ dB.

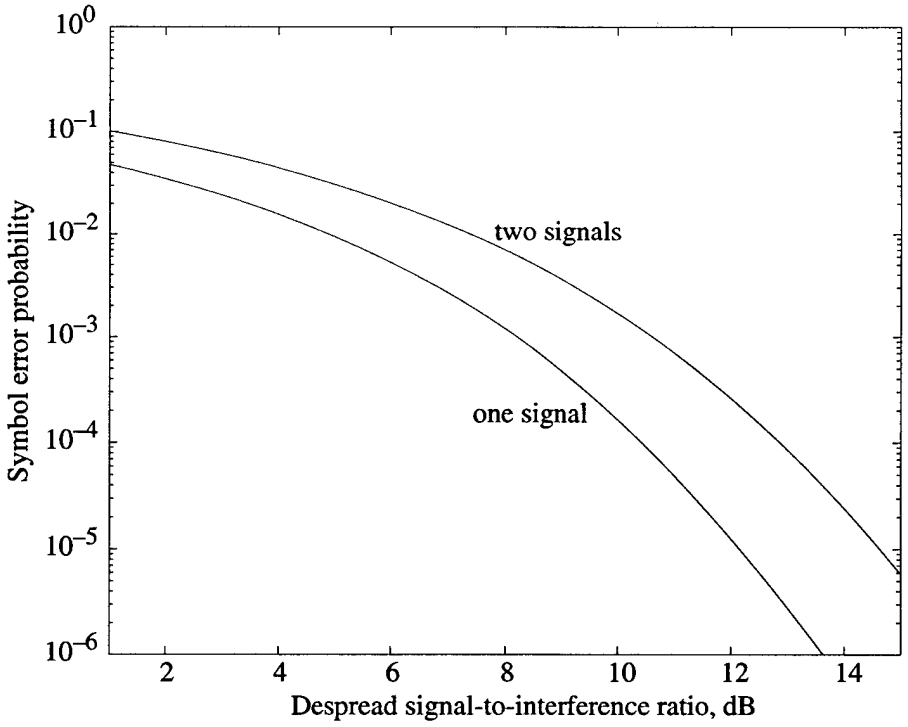


Figure 28. Symbol error probability of direct-sequence system with PSK in presence of multiple-access interference signal and $E_s/N_0 = 15$ dB.

The preceding bounding methods can be extended to the components of ϕ in $P_s(\phi)$ by observing that $\cos^2 \phi_i \leq 1$ and setting $X = \cos^2 \phi_i$ during the successive applications of Jensen's inequality, which is applicable if (3-51) is satisfied. After applying (3-42), we obtain

$$Q\left(\sqrt{\frac{2E_s}{N_{0l}}}\right) \leq P_s \leq Q\left(\sqrt{\frac{2E_s}{N_{0u}}}\right) \tag{3-56}$$

where

$$N_{0l} = N_0 + h\frac{I_t}{G}T_s \tag{3-57}$$

$$N_{0u} = N_0 + 2\frac{I_t}{G}T_s \tag{3-58}$$

$$I_t = \sum_{i=1}^{K} I_i \tag{3-59}$$

An approximation is provided by

$$P_s \approx Q\left(\sqrt{\frac{2E_s}{N_0 + \sqrt{2h}\, I_t T_s/G}}\right) \tag{3-60}$$

If P_s is specified, then the error in the required E_s/I_t caused by using (3-60) instead of (3-40) is bounded by $10 \log_{10}\sqrt{2/h}$ in decibels. Thus, the error is bounded by 2.39 dB for rectangular pulses and 2.66 dB for sinusoidal pulses.

The lower bound in (3-56) gives the same result as that often called the *standard Gaussian approximation*, in which V_1 in (3-24) is assumed to be approximately Gaussian, each ϕ_i in (3-30) is assumed to be uniformly distributed over $[0, 2\pi)$, and each τ_i is assumed to be uniformly distributed over $[0, T_c)$. This approximation, gives an optimistic result for P_s that can be as much as 4.77 dB in error for rectangular pulses according to (3-56). The substantial improvement in accuracy provided by (3-54) and (3-42) is due to the application of the Gaussian approximation only after conditioning V_1 on given values of ϕ and τ.

Figure 29 illustrates the symbol error probability for $K = 1$ as a function of GS/I for the standard Gaussian approximation of (3-56), the rough approximation of (3-60), and the upper and lower bounds given by (3-44), (3-52), and (3-42). The large error in the standard Gaussian approximation is evident. Although the rough approximation is reasonably accurate if $P_s \leq 10^{-2}$ in this example, it is increasingly less accurate as K increases.

For synchronous networks, (3-37) and (3-38) can be simplified because the $\{\tau_i\}$ are all zero. For either rectangular or sinusoidal pulses, we obtain

$$P_s(\phi) = Q\left(\sqrt{\frac{2E_s}{N_{0e}(\phi)}}\right) \tag{3-61}$$

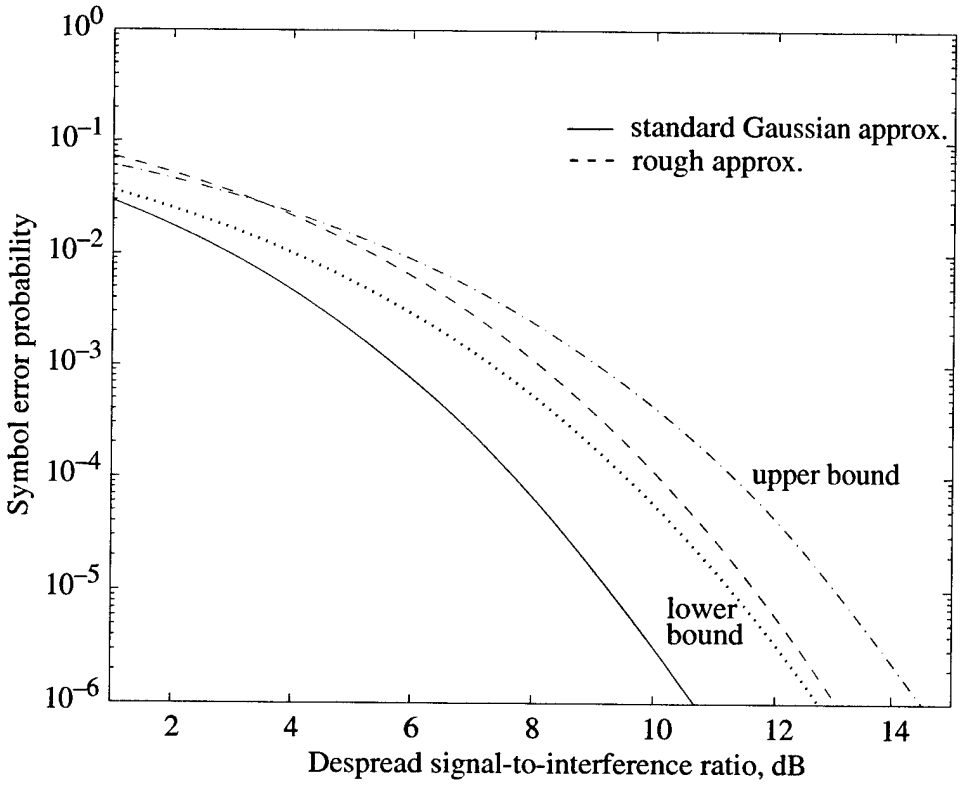


Figure 29. Symbol error probability of direct-sequence system with PSK in presence of single multiple-access interference signal and $E_s/N_0 = 15$ dB.

where

$$N_{0e}(\phi) = N_0 + \sum_{i=1}^{K} 2\frac{I_i}{G} T_s \cos^2 \phi_i \tag{3-62}$$

A comparison with (3-44) and (3-45) indicates that P_s for a synchronous network equals or exceeds P_s for a similar asynchronous network when random spreading sequences are used. This phenomenon is due to the increased bandwidth of a despread asynchronous interference signal, which allows increased filtering in the receiver.

3.4 Quadriphase Direct-Sequence Systems with Random Sequences

A general *quaternary direct-sequence signal* can be represented by

$$s(t) = \sqrt{S} m_1(t) p_1(t) \cos 2\pi f_0 t + \sqrt{S} m_2(t - t_0) p_2(t - t_0) \sin 2\pi f_0 t \tag{3-63}$$

where two spreading waveforms, $p_1(t)$ and $p_2(t)$, and two data signals, $m_1(t)$ and $m_2(t)$, are used with two quadrature carriers, and t_0 is the relative delay between the in-phase and quadrature components of the signal. For a *quadriphase direct-sequence system*, which uses quadriphase-shift keying (QPSK), $t_0 = 0$. For a direct-sequence system with offset QPSK or minimum-shift keying (MSK), $|t_0| = T_c/2$. For *offset* QPSK, the chip waveforms are rectangular; for MSK, they are sinusoidal. The use of MSK limits the spectral sidelobes of the direct-sequence signal, which may interfere with other signals.

Let T_s denote the duration of the data symbols before the generation of (3-63), and let $T_{s1} = 2T_s$ denote the duration of the transmitted channel symbols of $m_1(t)$ and $m_2(t)$. Let T_c denote the common chip duration of $p_1(t)$ and $p_2(t)$. It is assumed that the synchronization is perfect in the receiver, which is shown in Figure 30. Consequently, if the received signal is given by (3-5), then the sampled output of the upper integrator at the end of a symbol interval is given by

$$V = \pm\sqrt{S}\,T_{s1} + V_1 + V_2 \tag{3-64}$$

where

$$V_1 = \int_0^{T_{s1}} 2i(t)p_1(t)\cos 2\pi f_0 t\,dt \tag{3-65}$$

$$V_2 = \int_0^{T_{s1}} 2n(t)p_1(t)\cos 2\pi f_0 t\,dt \tag{3-66}$$

The term representing crosstalk,

$$V_3 = \int_0^{T_{s1}} \sqrt{S}m_2(t-t_0)p_2(t-t_0)p_1(t)\sin 4\pi f_0 t\,dt \tag{3-67}$$

is negligible if $f_0 >> 1/T_c$ so that the sinusoid in (3-67) varies much more rapidly than the product of the spreading waveforms. Similarly, the output of the lower integrator at the end of a channel-symbol interval is

$$U = \pm\sqrt{S}\,T_{s1} + U_1 + U_2 \tag{3-68}$$

where

$$U_1 = \int_{t_0}^{t_0+T_{s1}} 2i(t)p_2(t-t_0)\sin 2\pi f_0 t\,dt \tag{3-69}$$

$$U_2 = \int_{t_0}^{t_0+T_{s1}} 2n(t)p_2(t-t_0)\sin 2\pi f_0 t\,dt \tag{3-70}$$

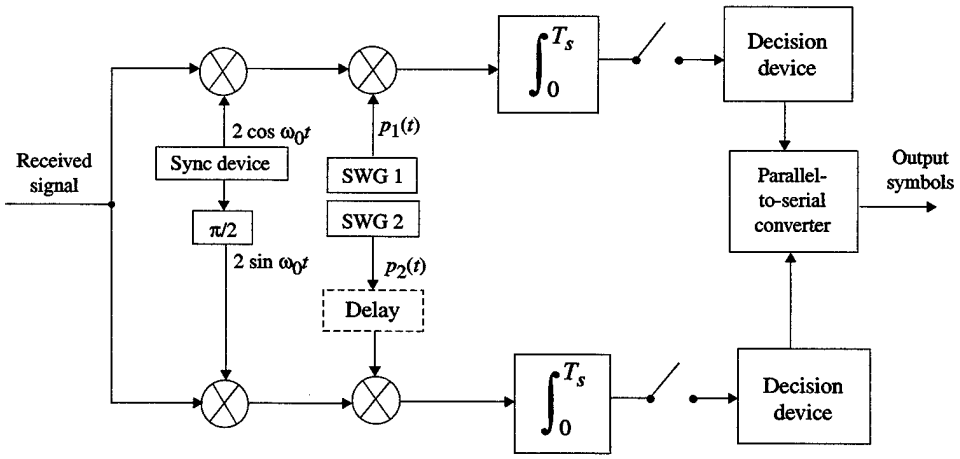


Figure 30. Receiver for direct-sequence signal with classical quaternary modulation (delay = 0 for QPSK and delay = $T_c/2$ for OQPSK and MSK); SWG = spreading waveform generator.

An analysis similar to the preceding one for a direct-sequence system with PSK indicates that the probability of an error is the same for both inputs of the parallel-to-serial converter and, hence, for each binary symbol of the data output. Of the available desired-signal power S, half is in each of the two components. Since $T_{s1} = 2T_s$, the energy per channel symbol is $E_s = ST_s$, the same as for a direct-sequence system with PSK, and

$$E[V] = E[U] = \pm\sqrt{S}T_{s1} = \pm 2\sqrt{S}T_s \tag{3-71}$$

Consider a network of classical quadriphase direct-sequence systems with $t_0 = 0$, each of which uses a pair of independent, random spreading sequences. A suitable model for multiple-access interference is

$$i(t) = \sum_{i=1}^{K} [\sqrt{I}\, q_{1i}(t-\tau_i)\cos(2\pi f_0 t + \phi_i) + \sqrt{I}\, q_{2i}(t-\tau_i)\sin(2\pi f_0 t + \phi_i)] \tag{3-72}$$

where $q_{1i}(t)$ and $q_{2i}(t)$ both have the form of (3-27) and incorporate the data modulation. A straightforward calculation using (3-65) indicates that V_1 is given by (3-24) with $G = T_{s1}/T_c$ and

$$\begin{aligned} J_\nu \;=\; & \sum_{i=1}^{K} \sqrt{I_i}\{\cos\phi_i[q_{1i,\nu-1}R_\psi(\tau_i) + q_{1i,\nu}R_\psi(T_c - \tau_i)] \\ & - \sin\phi_i[q_{2i,\nu-1}R_\psi(\tau_i) + q_{2i,\nu}R_\psi(T_c - \tau_i)]\} \end{aligned} \tag{3-73}$$

The statistical independence of the two sequences, the lemma of Section 3.3 and analogous results for U_1 yield

$$\mathrm{var}(V_1) = \mathrm{var}(U_1) = \sum_{i=1}^{K} \frac{T_{s1}}{T_c} I_i [R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i)] \tag{3-74}$$

A derivation similar to the one leading to (3-11) gives

$$\mathrm{var}(V_2) = \mathrm{var}(U_2) = N_0 T_{s1} \tag{3-75}$$

Since all variances are independent of ϕ, the Gaussian approximation yields a $P_s(\phi, \tau)$ that is independent of ϕ. Therefore,

$$P_s = \left(\frac{1}{T_c}\right)^K \int_0^{T_c} \cdots \int_0^{T_c} Q\left(\sqrt{\frac{2E_s}{N_{0e}(\tau)}}\right) d\tau \tag{3-76}$$

where

$$N_{0e}(\tau) = N_0 + \sum_{i=1}^{K} \frac{I_i}{T_c} [R_\psi^2(\tau_i) + R_\psi^2(T_c - \tau_i)] \tag{3-77}$$

In another version of the quaternary direct-sequence system, the same data symbols are carried by the in-phase and quadrature components, which implies that the received direct-sequence signal has the form given by (3-63)

with $m_1(t) = m_2(t) = m(t)$. Thus, although the spreading is done by quadrature carriers, the data modulation may be regarded as binary PSK. A receiver for this *balanced quaternary system* is shown in Figure 31. The synchronization system, which is assumed to operate perfectly in the subsequent analysis, is not shown for simplicity. If $f_0 >> 1/T_c$, the crosstalk terms similar to (3-67) are negligible, and the input to the decision device is

$$V = \pm 2\sqrt{S}\,T_s + V_1 + V_2 + U_1 + U_2 \tag{3-78}$$

where $T_{s1} = T_s$ is the duration of both a data symbol and a channel symbol. If $p_1(t)$ and $p_2(t)$ can be approximated by independent random binary sequences, then V_1, V_2, U_1, and U_2 are zero-mean uncorrelated random variables. Therefore, the variance of V is equal to the sum of the variances of these four random variables, and

$$E[V] = \pm 2\sqrt{S}T_s \tag{3-79}$$

An analysis similar to that for classical quadriphase signals yields (3-76) again. Thus, the classical and balanced quadriphase systems perform equally well against multiple-access interference.

Using the bounding and approximation methods previously developed, we find that the symbol error probability of (3-76) satisfies

$$Q\left(\sqrt{\frac{2E_s}{N_{0l}}}\right) \le P_s \le Q\left(\sqrt{\frac{2E_s}{N_{0u}}}\right) \tag{3-80}$$

where

$$N_{0l} = N_0 + h\frac{I_t}{G}T_s \tag{3-81}$$

$$N_{0u} = N_0 + \frac{I_t}{G}T_s \tag{3-82}$$

and the total interference power I_t is defined by (3-59). The convexity required to derive the lower bound is valid if

$$E_s \ge \frac{3}{2}\left(N_0 + \frac{I_t}{G}T_s\right) \tag{3-83}$$

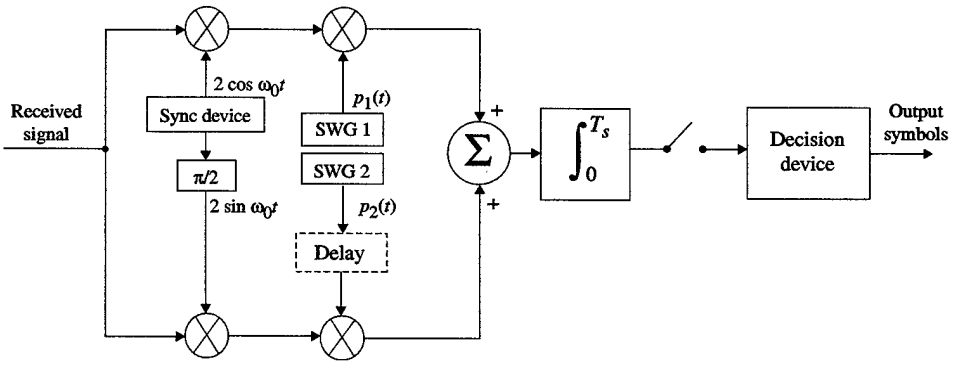


Figure 31. Receiver for direct-sequence signal with balanced quaternary modulation (delay = 0 for QPSK and delay = $T_c/2$ for OQPSK and MSK); SWG = spreading waveform generator.

An approximation that limits the error in the required E_s/I_t for a specified P_s to 10 $\log_{10}(1/\sqrt{h})$ is

$$P_s \approx Q\left(\frac{2E_s}{N_0 + \sqrt{h}I_tT_s/G}\right) \tag{3-84}$$

This approximation introduces errors bounded by 0.88 dB and 1.16 dB for rectangular and sinusoidal pulses, respectively, which are much less than the maximum errors for (3-60). In (3-80) and (3-84), only the total interference power is relevant, not how it is distributed among the individual interference signals.

For synchronous networks with either rectangular or sinusoidal pulses, we set the $\{\tau_i\}$ equal to zero in (3-77) and obtain

$$P_s = Q\left(\frac{2E_s}{N_0 + I_tT_s/G}\right) \tag{3-85}$$

Since this equation coincides with the upper bound in (3-80), we conclude that asynchronous networks accommodate more multiple-access interference than similar synchronous networks using quadriphase direct-sequence signals with random spreading sequences.

To compare asynchronous quadriphase direct-sequence systems with asynchronous systems using PSK, we find a lower bound on P_s for direct-sequence systems with PSK. Substituting (3-37) into (3-40) and applying Jensen's inequality successively to the integrations over $\phi_i, i = 1, 2, \ldots, K$, we find that a lower bound on P_s is given by the right-hand side of (3-76) if (3-83) is satisfied. This result implies that asynchronous quadriphase direct-sequence systems are more resistant to multiple-access interference than asynchronous direct-sequence systems with PSK.

3.5 Cellular Networks and Power Control

In a *cellular network*, a geographic region is partitioned into cells, as illustrated in Figure 32. A base station that includes a transmitter and receiver is located at the center of each cell. Ideally, the cells have equal hexagonal areas. Each *mobile* (user or subscriber) in the network transmits omnidirectionally and communicates with the base station from which it receives the largest average power. Typically, most of the mobiles in a cell communicate with the base station at the center of the cell, and only a few communicate with more distant ones. The base stations act as switching centers for the mobiles and communicate among themselves by wirelines in most applications. Cellular networks with direct-sequence CDMA allow universal frequency reuse in that the same carrier frequency and spectral band is shared by all the cells. Distinctions among the direct-sequence signals are possible because each signal is assigned a unique spreading sequence.

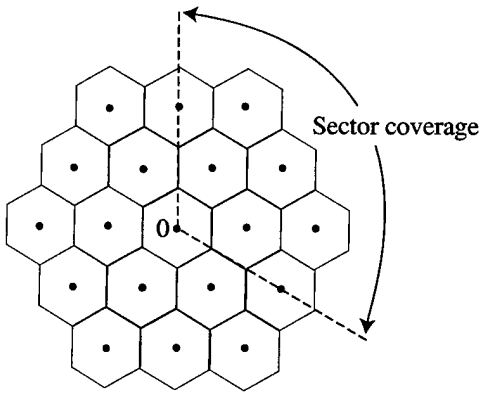


Figure 32. Geometry of cellular network with base station at center of each hexagon. Two concentric tiers of cells surrounding a central cell are shown.

Cells may be divided into *sectors* by using several directional sector antennas or arrays at the base stations. Only mobiles in the directions covered by a sector antenna can cause multiple-access interference on the *reverse link* or *uplink* from a mobile to its associated sector antenna. Only a sector antenna serving a cell sector oriented toward a mobile can cause multiple-access interference on the *forward link* or *downlink* from the mobile's associated sector antenna to the mobile. Thus, the numbers of interfering signals on both the uplink and the downlink are reduced approximately by a factor equal to the number of sectors.

The principal difficulty of direct-sequence CDMA is called the *near-far problem*. If all mobiles transmit at the same power level, then the received power at a base station is higher for transmitters near the receiving antenna. There is a near-far problem because transmitters that are far from the receiving antenna may be at a substantial power disadvantage, and the spread-spectrum processing gain may not be enough to allow satisfactory reception of their signals. A similar problem may also result from large differences in received power levels due to differences in the shadowing experienced by signals traversing different paths or due to independent fading.

In cellular communication networks, the near-far problem is critical only on the uplink because on the downlink, the base station transmits orthogonal signals synchronously to each mobile associated with it. For cellular networks, the usual solution to the near-far problem of uplinks is *power control*, whereby all mobiles regulate their power levels. By this means, power control potentially ensures that the power arriving at a common receiving antenna is almost the same for all transmitters. Since solving the near-far problem is essential to the viability of a direct-sequence CDMA network, the accuracy of the power control is a crucial issue.

In networks with *peer-to-peer communications*, there is no cellular or hierarchical structure. Communications between two mobiles are either direct or are relayed by other mobiles. Since there is no feasible method of power control to prevent the near-far problem, direct-sequence CDMA systems are not as attractive an option as frequency-hopping CDMA systems in these networks.

An *open-loop method* of power control in a cellular network causes a mobile to adjust its transmitted power to be inversely proportional to the received power of a *pilot signal* transmitted by the base station. An open-loop method is used to initiate power control, but its subsequent effectiveness requires that the propagation losses on the forward and reverse links be nearly the same. Whether they are or not depends on the duplexing method used to allow simultaneous or nearly simultaneous transmissions on both links. *Frequency-division duplexing* assigns different frequencies to an uplink and its corresponding downlink. *Time-division duplexing* assigns closely spaced but distinct time slots to the two links. When frequency-division duplexing is used, as in the IS-95 and Global System for Mobile (GSM) standards, the frequency separation is generally wide enough that the channel transfer functions of the uplink and downlink are different. This lack of *link reciprocity* implies that power measurements over the downlink do not provide reliable information for subsequent uplink transmissions. When time-division duplexing is used, the received local-mean power levels for the uplink and the downlink will usually be nearly equal when the transmitted powers are the same, but the Rayleigh fading may subvert link reciprocity. For these reasons, a *closed-loop method* of power control, which is more flexible than an open-loop method, is desirable. A closed-loop method requires the base station to transmit power-control information to each mobile based on the power level received from the mobile or the signal-to-interference power ratio.

When closed-loop power control is used, each base station attempts to either directly or indirectly track the received power of a desired signal from a mobile and dynamically transmit a power-control signal [22], [23]. The effect of increasing the carrier frequency or the mobile speeds is to increase the fading rate. As the fading rate increases, the tracking ability and, hence, the power-control accuracy decline. This problem is often dismissed by invoking the putative tradeoff between the power control and the bit or symbol interleaving [22], [23]. It is asserted that the large fade durations during slow fading enable effective power control, whereas the imperfect power control in the presence of fast fading is compensated by the increased time diversity provided by the interleaving and channel coding. However, this argument ignores both the potential severity of the near-far problem and the limits of compensation as the fading rate increases. If the power control breaks down completely, then close interfering mobiles can cause frequent error bursts of duration long enough to overwhelm the ability of the deinterleaver to disperse the errors so that the decoder can eliminate them. Thus, some degree of power control must be maintained as the vehicle speeds or the carrier frequency increases. The degree required when the interleaving is perfect is quantified subsequently.

The following performance analysis of the uplink [24] begins with the derivation of the intercell interference factor, which is the ratio of the intercell interference power to the intracell interference power. The *intercell interference* arrives from mobiles associated with different base stations than the one receiving a desired signal. The *intracell interference* arrives from mo-

biles that are associated with the same base station receiving a desired signal. The performance is evaluated using two different criteria: the SINR and the bit error rate. The SINR criterion has the advantage that it simplifies the analysis and does not require specification of the data modulation or channel coding. The bit-error-rate criterion has the advantage that the impact of the channel coding can be calculated. For both criteria, the fading is flat and no explicit diversity or rake combining is assumed. Since the interference signals arrive asynchronously, they cannot be suppressed by using orthogonal spreading sequences.

3.6 Intercell Interference of Uplink

To account for the fading and instantaneous power control in a mathematically tractable way, the shadowing and fading factors in (1-4) are approximated [1] by a lognormal random variable. Thus, at a particular time it is assumed that the *equivalent shadowing factor* η implicitly defined by

$$10^{\eta/10} = 10^{\xi/10}\alpha^2 \tag{3-86}$$

has a probability density function that is approximately Gaussian. This equation, the statistical independence of ξ and α, and the fact that $E[\xi] = 0$ imply that

$$E[\eta] = \frac{2}{b}E[ln\,\alpha] \tag{3-87}$$

$$E[\eta^2] = E[\xi^2] + \frac{4}{b^2}E[(ln\,\alpha)^2] \tag{3-88}$$

where $b = (ln\,10)/10$. To evaluate these equations when α has the Nakagami-m density function of (1-29), we express the expectations as integrals, change the integration variables, and apply the identities [8]

$$\int_0^\infty x^{\nu-1}e^{-\mu x}ln\,x\,dx = \frac{\Gamma(\nu)}{\mu^\nu}[\psi(\nu) - ln\,\mu] \tag{3-89}$$

$$\int_0^\infty x^{\nu-1}e^{-\mu x}(ln\,x)^2\,dx = \frac{\Gamma(\nu)}{\mu^\nu}\{[\psi(\nu) - ln\,\mu]^2 + \zeta(2,\nu)\} \tag{3-90}$$

where $\mathrm{Re}(\mu) > 0$, $\mathrm{Re}(\nu) > 0$, $\psi(\nu)$ is the psi function given by

$$\psi(\nu) = \sum_{i=1}^{\nu-1}\frac{1}{i} - C\,, \quad C \cong 0.5772 \tag{3-91}$$

when ν is a positive integer, and $\zeta(2,\nu)$ is the Riemann zeta function given by

$$\zeta(2,\nu) = \sum_{i=0}^{\infty}\frac{1}{(\nu+i)^2}\,, \quad \nu \neq 0, -1, -2, \ldots \tag{3-92}$$

Let σ_η^2 denote the variance of η. Since $E[\xi^2] = \sigma_s^2$, the variance of ξ, we find that

$$E[\eta] = \frac{1}{b}[\psi(m) - ln(m)] \tag{3-93}$$

$$\sigma_\eta^2 = \sigma_s^2 + \frac{\zeta(2,m)}{b^2} \tag{3-94}$$

The impact of the fading declines with increasing m. For Rayleigh fading, $m = 1$ and $\zeta(2, 1) = 1.65$, so $E[\eta] = -2.5$ and $\sigma_\eta^2 = \sigma_s^2 + 31.0$. For $m = 5$, which approximates Ricean fading with Rice factor $\kappa = 8.47$, $E[\eta] = -0.45$ and $\sigma_\eta^2 = \sigma_s^2 + 4.2$.

Consider a cellular network in which each base station is located at the center of a hexagonal area, as illustrated in Figure 32. To analyze uplink interference, it is assumed that the desired signal arrives at base station 0, while the other base stations are labeled 1, 2, ..., N_B. The directions covered by one of three sectors associated with base station 0 are indicated in the figure. Each mobile in the network transmits omnidirectionally and is associated with the base station from which it receives the largest average short-term or *instantaneous power*. This base station establishes the uplink power control of the mobile. If a mobile is associated with base station i, then (1-1), (1-4), and (3-86) indicate that the instantaneous power received by base station j is

$$D_{ij} = p_{0i}\left(\frac{r_j}{R_0}\right)^{-\beta} 10^{\eta_j/10} = p_{0i}\left(\frac{r_j}{R_0}\right)^{-\beta} \exp(b\,\eta_j) \tag{3-95}$$

where r_j is the distance to base station j, η_j is the equivalent shadowing factor, p_{0i} is the area-mean power at $r_j = R_0$, and it is assumed that the attenuation power-law β is the same throughout the network. If the power control exerted by base station i ensures that it receives unit instantaneous power from each mobile associated with it, then $D_{ii} = 1$. Consequently, $p_{0i} = (r_i/R_0)^\beta \exp(-b\,\eta_i)$, and

$$D_{ij} = \left(\frac{r_i}{r_j}\right)^\beta \exp[b(\eta_j - \eta_i)] \tag{3-96}$$

Assuming a common fading model for all of the $\{\eta_i\}$, (3-87) implies that they all have the same mean value. The form of (3-96) then indicates that this common mean value is irrelevant to the statistics of D_{ij} and hence can be ignored without penalty in the subsequent statistical analysis of D_{ij}. The simplifying approximation is made that the base station with which a mobile is associated receives more instantaneous power than any other station, and hence $D_{ij} \leq 1$. This inequality is exact if the propagation losses on the uplink and downlink are the same.

The probability distribution function of the interference power D_{i0} at base station 0 given that the mobile producing the interference is associated with base station i is [25]

$$F_i(x) = Pr[D_{i0} \leq x \mid D_{ij} \leq 1, 0 \leq j \leq N_B] = \frac{\phi_i(x)}{\phi_i(1)} \tag{3-97}$$

where

$$\phi_i(x) = Pr[D_{i0} \leq x; D_{ij} \leq 1, 0 \leq j \leq N_B] \,. \tag{3-98}$$

Thus, $F_i(x) = 0$ if $x < 0$, and $F_i(x) = 1$ if $x \geq 1$. Let

$$\phi_i(x \mid \eta_i, r_i, \theta_i) = Pr[D_{i0} \leq x; D_{ij} \leq 1, 0 \leq j \leq N_B \mid \eta_i, r_i, \theta_i] \tag{3-99}$$

where this probability is conditional on η_i, the equivalent shadowing factor for the controlling base station, and the polar coordinates r_i, θ_i of the mobile relative to base station i. It is assumed that each of the $\{\eta_j\}$ is statistically independent with the common variance σ_η^2. Therefore, given η_i, D_{ij} and $D_{ik}, j \neq k$, are statistically independent. Since each of the $\{\eta_j\}$ has a Gaussian probability density function, (3-96) implies that for $0 \leq x \leq 1$,

$$\phi_i(x \mid \eta_i, r_i, \theta_i) = Q_c\left(\frac{b\eta_i + \beta\, ln(r_0/r_i) + ln\, x}{b\sigma_\eta}\right) \prod_{\substack{j=1 \\ j\neq i}}^{N_B} Q_c\left(\frac{b\eta_i + \beta\, ln(r_j/r_i)}{b\sigma_\eta}\right) \tag{3-100}$$

where $Q_c(x) = 1 - Q(x)$, $Q(x)$ is defined by (2-54), and $r_j, 0 \leq j \leq N_B$, is a function of r_i, θ_i, and the location of base station j.

The probability $\phi_i(x)$, and hence the distribution $F_i(x)$, can be determined by evaluating the expected value of (3-100) with respect to the random variables η_i, r_i, and θ_i. If a mobile is associated with base station i, then its location is assumed to be uniformly distributed within a circle of radius R_b surrounding the base station. Therefore,

$$\phi_i(x) = \int_0^{2\pi} d\theta \int_0^{R_b} dr \int_{-\infty}^{\infty} d\eta \, \frac{r \exp\left(-\frac{\eta^2}{2\sigma_\eta^2}\right)}{\sqrt{2\pi}\,\sigma_\eta\,\pi R_b^2} \phi_i(x \mid \eta, r, \theta) \tag{3-101}$$

which determines the distribution function in (3-97).

Let I_{te} denote the total intercell interference relative to the unit desired-signal power that each base station attempts to maintain by power control. Let K denote the number of active mobiles associated with a base station or sector antenna, which may be a random variable because of voice-activity detection or the movement of mobiles among the cells. Since $E[D_{i0}]$ and $\text{var}[D_{i0}]$ are the same for all mobiles associated with base station i, a straightforward calculation yields

$$E[I_{te}] = E[K] \sum_{i=1}^{N_B} E[D_{i0}] \tag{3-102}$$

$$\text{var}[I_{te}] = E[K] \sum_{i=1}^{N_B} \text{var}[D_{i0}] + \text{var}[K] \left(\sum_{i=1}^{N_B} E[D_{i0}]\right)^2 \tag{3-103}$$

In general, $E[I_{te}]$ and $\text{var}[I_{te}]$ decrease as the attenuation power law β increases. The *intercell interference factor*, $g = E[I_{te}]/E[K]$, is the ratio of the

average intercell interference power to the average intracell interference power. Table 1, calculated in [25], lists g versus σ_η when $N_B = 60$ cells in four concentric tiers surrounding a central cell, R_b is five times the distance from a base station to the corner of its surrounding hexagonal cell, and $\beta = 4$. The dependence of g on the specific fading model is exerted through (3-94), which relates σ_η to m and σ_s. Table 1 also lists the *variance factor* $g_1 = \text{var}[I_{te}]/E[K]$ assuming that $\text{var}[K] = 0$.

The results in Table 1 depend on the pessimistic assumption that the equivalent shadowing factors from a mobile to two different base stations are independent random variables. Suppose, instead, that each factor is the sum of a common component and an equal-power independent component that depends on the receiving base station. Then (3-96) implies that the common component cancels. As a result, in determining g from Table 1, the effective value of σ_η is reduced by a factor of $\sqrt{2}$ relative to what it would be without the common component.

Since $\sigma_\eta^2 = \sigma_s^2 + 31.0$ for Rayleigh fading and Table 1 indicates that g increases slowly with σ_η, the effect of the fading is unimportant or negligible if $\sigma_s \geq 6$ dB, which is usually satisfied in practical networks. If it is assumed, as is tacitly done by many authors, that the power control is based on a long-term-average power estimate that averages out the fading, then the preceding equations and Table 1 are valid with $\sigma_\eta = \sigma_s$.

3.7 SINR Analysis of Conditional Outage Probability

For a direct-sequence CDMA system, it is assumed that the total power I_t of the multiple-access interference is approximately uniformly distributed over the signal bandwidth, which is approximately equal to $1/T_c$. For instantaneous power control, the instantaneous SINR is defined to be $E_s/(N_0 + I_t T_c)$, the ratio of the received energy per symbol E_s to the equivalent power spectral density of the interference plus noise. An *outage* is said to occur if the instantaneous SINR is less than a specified threshold Z, which may be adjusted to account for any diversity or rake combining. In this section, the interference is assumed to arise from $K - 1$ other active mobiles in a single cell or sector. Let $E_i = I_i T_s$, $i = 1, 2, \ldots, K - 1$, denote

Table 1. Interference factor and variance factor as functions of σ_η when $\text{var}[K] = 0$.

σ_η, dB	g	g_1
3	0.460	0.137
4	0.486	0.143
$6/\sqrt{2}$	0.493	0.145
5	0.519	0.153
$8/\sqrt{2}$	0.544	0.162
6	0.558	0.167
7	0.598	0.183
$10/\sqrt{2}$	0.601	0.184
8	0.634	0.189

the received energy in a symbol due to interference signal i with power I_i. These definitions imply that an outage occurs if

$$E_s Z^{-1} < N_0 + \frac{1}{G} \sum_{i=1}^{K-1} E_i \tag{3-104}$$

where $G = T_s/T_c$ is the processing gain. Let E_{s0} denote the common desired energy per symbol for all the signals associated with the base station of a single cell sector. When instantaneous power control is used, $E_s = E_{s0}\,\epsilon_0$ and $E_i = E_{s0}\,\epsilon_i$, $i = 1, 2, \ldots, K-1$, where ϵ_0 and ϵ_i are random variables that account for imperfections in the power control. Substitution into (3-104) yields the outage condition

$$G(Z^{-1}\epsilon_0 - \gamma_0^{-1}) < X \tag{3-105}$$

where $\gamma_0 = E_{s0}/N_0$ is the energy-to-noise density ratio of the desired signal when the power control is perfect, and we define

$$X = \sum_{i=1}^{K-1} \epsilon_i \tag{3-106}$$

By analogy with the lognormal spatial variation of the local-mean power, each of the $\{\epsilon_i\}$ is modeled as an independent lognormal random variable. Therefore,

$$\epsilon_i = 10^{\xi_i/10} = \exp(b\xi_i), \quad i = 0, 1, 2, \ldots, K-1 \tag{3-107}$$

where each of the $\{\xi_i\}$ is a zero-mean Gaussian random variable with common variance σ_e^2. The moments of ϵ_i can be derived by direct integration or from the moment-generating function of ξ_i. We obtain

$$E[\epsilon_i] = \exp\left(\frac{b^2\sigma_e^2}{2}\right), \quad E[\epsilon_i^2] = \exp(2b^2\sigma_e^2) \tag{3-108}$$

If K is a constant, then the mean $\bar{X}$ and the variance σ_x^2 of X in (3-106) are

$$\bar{X} = (K-1)\exp\left(\frac{b^2\sigma_e^2}{2}\right), \; \sigma_x^2 = (K-1)[\exp(2b^2\sigma_e^2) - \exp(b^2\sigma_e^2)] \tag{3-109}$$

The random variable X is the sum of $K-1$ lognormally distributed random variables. Since the distribution of X cannot be compactly expressed in closed form when $K > 3$, two approximate methods are adopted. The first method is based on the central limit theorem, and the second method is based on the assumption that σ_e is small.

Since X is the sum of $K-1$ independent, identically distributed random variables, each with a finite mean and variance, the central limit theorem implies that the probability distribution function of X is approximately Gaussian when K is sufficiently large. Consequently, given the values of K and ϵ_0, the conditional probability of outage may be calculated from (3-105). Using (3-107) and integrating over the Gaussian density function of ξ_0, we then obtain the conditional probability of outage given the value of $K >> 1$:

$$P_{\text{out}}(K) = \int_{-\infty}^{\infty} Q\left[\frac{G(Z^{-1}e^{b\xi} - \gamma_0^{-1}) - \bar{X}}{\sigma_x}\right] \frac{\exp(-\xi^2/2\sigma_e^2)}{\sqrt{2\pi}\sigma_e}\, d\xi \qquad (3\text{-}110)$$

As $\sigma_e \to 0$ and hence $\sigma_x \to 0$, $P_{\text{out}}(K)$ approaches a step function.

In the second approximate method, it is assumed that σ_e is sufficiently small and K is sufficiently large that $\sigma_x << \bar{X}$. From (3-109), it is observed that a sufficient condition for this assumption is that

$$\sqrt{K-1} >> \exp\left(\frac{b^2\sigma_e^2}{2}\right) \qquad (3\text{-}111)$$

The assumption implies that X is well approximated by the constant $\bar{X}$ given by (3-109). Since the only remaining random variable in (3-105) is $\epsilon_0 = \exp(b\xi_0)$, it follows that

$$P_{\text{out}}(K) = Q\left\{-\frac{ln[(K-1)G^{-1}Z\exp(b^2\sigma_e^2/2) + Z\gamma_0^{-1}]}{b\sigma_e}\right\} \qquad (3\text{-}112)$$

3.8 Variations in the Number of Active Mobiles

In the derivations of (3-110) and (3-112), the number of mobiles actively transmitting, K, is held constant. However, it is appropriate to model K as a random variable because of the movement of mobiles into and out of each sector and the changing of the cell or sector antenna with which a mobile communicates. Furthermore, a potentially active mobile may not be transmitting; for voice communications with voice-activity detection, energy transmission typically is necessary only roughly 40% of the time. As is shown below, a discrete random variable K with a Poisson distribution incorporates both of these effects.

To simplify the analysis, it is assumed that the average number of mobiles associated with each cell or sector antenna is the same and that the location of a mobile is uniformly distributed throughout a region. Let q denote the probability that a potentially transmitting mobile is actively transmitting. Then the probability that an active mobile is associated with a particular cell or sector antenna is $\mu q/N_r$, where N_r is the number of mobiles in the region and μ is the average number of mobiles per sector. If the N_r mobiles are independently located in the region, then the probability of $K = k$ active mobiles being associated with a sector antenna is given by the binomial distribution

$$P(N_r, k) = \binom{N_r}{k}\left(\frac{\lambda}{N_r}\right)^k\left(1 - \frac{\lambda}{N_r}\right)^{N_r-k} \qquad (3\text{-}113)$$

where $\lambda = \mu q$ is assumed to be a constant. This equation can be expressed as

$$P(N_r, k) = \frac{(1-1/N_r)(1-2/N_r)\ldots(1-(k-1)/N_r)}{k!}\lambda^k\left(1-\frac{\lambda}{N_r}\right)^{-k}\left(1-\frac{\lambda}{N_r}\right)^{N_r} \quad (3\text{-}114)$$

As $N_r \rightarrow \infty$, the initial fraction $\rightarrow 1/k!, (1-\lambda/N_r)^{-k} \rightarrow 1$, and $(1-\lambda/N_r)^{N_r} \rightarrow \exp(-\lambda)$. Therefore, $P(N_r, k)$ approaches

$$P_u(k) = \frac{\exp(-\lambda)\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots \quad (3\text{-}115)$$

which is the Poisson distribution function.

Since the desired mobile is assumed to be present, it is necessary to calculate the conditional probability that $K = k$ given that $K \geq 1$. From the definition of a conditional probability and (3-115), it follows that this probability is

$$P_c(k) = \frac{\exp(-\lambda)\lambda^k}{[1-\exp(-\lambda)]k!}, \quad k = 1, 2, \ldots \quad (3\text{-}116)$$

and $P_c(0) = 0$. Using this equation, the probability of outage is

$$P_{\text{out}} = \sum_{k=1}^{\infty} \frac{\exp(-\lambda)\lambda^k}{[1-\exp(-\lambda)]k!} P_{\text{out}}(k) \quad (3\text{-}117)$$

where $P_{\text{out}}(k)$ is given by (3-110) or (3-112).

The intercell interference from mobiles associated with other base stations introduces an additional average power equal to $g\mu q(E_{s0}/T_s)$ into a given base station, where g is the intercell interference factor. Accordingly, the impact of the intercell interference is modeled as equivalent to an average of $g\mu$ additional mobiles in a sector [26]. When intercell interference is taken into account, the equations of Section 3.7 for a single cell or sector are modified. The parameter μ is replaced by $\mu(1+g)$, and λ becomes the *equivalent number of mobiles*

$$\lambda = \mu q(1+g) \quad (3\text{-}118)$$

Figure 33 illustrates the probability of no outage, $1-P_{\text{out}}$, as a function of λ for various values of σ_e. Both approximate models, which give (3-110) and (3-112), are used in (3-117) to calculate the curves. Equations (3-110) and (3-112) indicate that the outage probability depends on the ratios G/Z and γ_0/G rather than on G, Z, and γ_0 separately. The parameter values for Figure 32 are $G/Z = 40$ and $\gamma_0/G = 0.5$, which could correspond to $Z = 7$ dB, $G = 23$ dB, and $\gamma_0 = 20$ dB. The closeness of the results for the two models indicates that when $\sigma_e \leq 2$ dB both models give accurate outage probabilities and the effect of power-control errors in the interference signals is unimportant. As an example of the application of the figure, suppose that

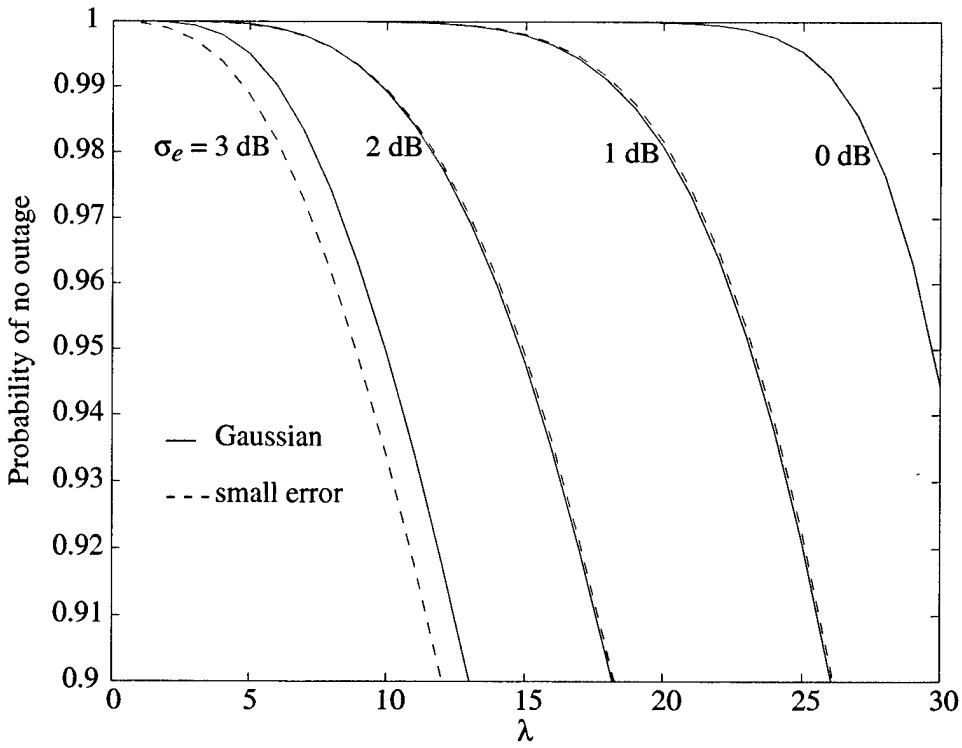


Figure 33. Probability of no outage for instantaneous power control, $G/Z = 40$, $\gamma_0/G = 0.5$, and $\sigma_e = 0, 1, 2, 3$ dB.

the attenuation power law is $\beta = 4$, $\sigma_\eta = 8$ dB, $\sigma_e = 1.0$ dB, and $1 - P_{\text{out}} = 0.95$ is desired. Table 1 gives $g = 0.63$. The figure indicates that $\lambda = 23$ is needed. If $q = 0.4$ due to voice-activity detection, the average number of mobiles per sector that can be accommodated is $\mu = 35.3$. For data communications, the network capacity is lower. For example, if $q = 1$, then the average number of mobiles per sector that can be accommodated is 14.1.

3.9 Local-Mean Power Control

When the instantaneous signal power cannot be tracked because of the fast multipath fading, one might consider measuring the local-mean power, which is a long-term-average power obtained by averaging out the fading component. This measurement enables the system to implement *local-mean power control.* Two different analyses of the effects of local-mean power control are presented.

In the first analysis, which explores the potential effectiveness of local-mean power control, all received signals experience Rayleigh fading and the local-mean power control is perfect. Therefore, the received energy levels are proportional to the squares of Rayleigh-distributed random variables and, hence, are exponentially distributed, as shown in Appendix A-4. Thus, $E_s = E_{s0}\,\epsilon_0$ and $E_i = E_{s0}\,\epsilon_i$, where each ϵ_i, $i = 0, 1, 2, \ldots, K-1$, is an independent random variable with the exponential probability density function:

$$f_s(x) = \exp(-x)u(x) \tag{3-119}$$

and E_{s0} is the desired value of the average energy per symbol after averaging over the fading. The probability distribution function of the sum of $K-1$ independent random variables, each with the exponential density of (3-119), is given by (A-50). Therefore, X in (3-106) has the distribution

$$F_X(x) = 1 - \exp(-x) \sum_{i=0}^{K-2} \frac{x^i}{i!}\ ,\ x \geq 0 \tag{3-120}$$

Conditioning on the value of ϵ_0, using (3-120) to evaluate the probability of the outage condition (3-105), and then removing the conditioning by using (3-119) yields

$$P_{\text{out}}(K) = \int_0^\infty e^{-\xi} \exp[-c(\xi)] \sum_{i=0}^{K-2} \frac{[c(\xi)]^i}{i!}\, d\xi \tag{3-121}$$

where

$$c(\xi) = GZ^{-1}\xi - G\gamma_0^{-1} \tag{3-122}$$

Replacing $[c(\xi)]^i$ by its binomial expansion, we obtain a double summation of integrals that can be evaluated using the gamma function defined by (A-12). After simplification, we obtain

$$P_{\text{out}}(K) = \sum_{i=0}^{K-2} \sum_{l=0}^{i} \frac{\exp(G\gamma_0^{-1})(GZ^{-1})^l(-G\gamma_0^{-1})^{i-l}}{(i-l)!(1+GZ^{-1})^{l+1}} \tag{3-123}$$

Interchanging the two sums and changing their limits accordingly, the inner sum is over a geometric series. Evaluating it, we obtain the final result:

$$P_{\text{out}}(K) = \exp(G\gamma_0^{-1}) \sum_{l=0}^{K-2} \frac{(-G\gamma_0^{-1})^l}{l!} \left[1 - \left(\frac{GZ^{-1}}{1+GZ^{-1}}\right)^{K-1-l}\right] \tag{3-124}$$

The probability of outage is determined by substitution into (3-117). When $\gamma_0 = \infty$, only the $l = 0$ term in (3-124) is nonzero. Substitution into (3-117) and evaluation of the sum yields

$$P_{\text{out}} = 1 - \frac{\exp\left(\frac{\lambda}{1+G^{-1}Z}\right) - 1}{\exp(\lambda) - 1}(1 + G^{-1}Z)\ ,\ \ \gamma_0 = \infty \tag{3-125}$$

For perfect local-mean power control and Rayleigh fading, (3-119) gives $E[\epsilon_i] = 1$ and $\text{var}(\epsilon_i) = 1$. Therefore, a sufficient condition for $\sigma_x << \bar{X}$ is that $\sqrt{K-1} >> 1$. If this condition is satisfied, then X is well approximated by $\bar{X} = K - 1$, which is equivalent to ignoring the fading of the multiple-access interference signals. With this approximation, the only remaining random variable in (3-105) is exponentially distributed, and hence the conditional probability of outage given K is

$$P_{\text{out}}(K) = 1 - \exp[-(K-1)G^{-1}Z - \gamma_0^{-1}Z] \tag{3-126}$$

Substituting this equation into (3-117) and evaluating the sum, we obtain the approximation

$$P_{\text{out}} = 1 - \frac{\exp(-\gamma_0^{-1}Z + G^{-1}Z)[\exp(\lambda \exp(-G^{-1}Z)) - 1]}{\exp(\lambda) - 1} \tag{3-127}$$

Figure 34 illustrates the probability of no outage as a function of λ for $G/Z = 40$ and two values of γ_0/G using either the approximation (3-127) or the more precise (3-125), (3-124), and (3-117). It is observed that neglecting the fading of the interference signals and using the approximation makes little difference in the results. The effect of $\gamma_0 = E_{s0}/N_0$ is considerable. A comparison of Figures 34 and 33 indicates that when Rayleigh fading occurs, even perfect local-mean power control is not as useful as imperfect instantaneous power control unless σ_e is very large.

Since accurate power measurements require a certain amount of time, whether a power-control scheme is instantaneous, local mean, or something intermediate depends on the fading rate. To reduce the fading rate so that the power control is instantaneous and accurate, one might minimize the carrier frequency or limit the size of cells if these options are available.

The second analysis of the effects of local-mean power control uses the preceding results to develop a simple approximation to previous performance calculations [26], [27]. This analysis has the advantages that the fading statistics do not have to be explicitly defined and the effect of imperfect local-mean power control is easily calculated. Let E_{sl} denote the local-mean

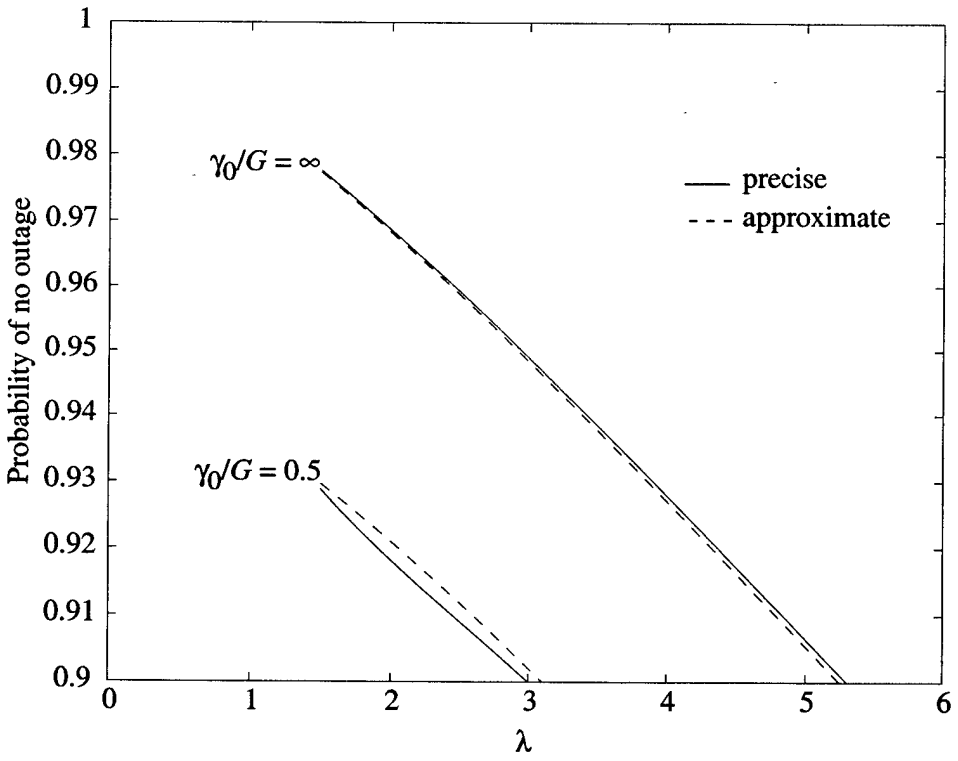


Figure 34. Probability of no outage for perfect local-mean power control, $G/Z = 40$, and $\gamma_0/G = 0.5, \infty$.

energy per symbol, which is defined as the average energy per symbol after averaging over the fading. Similarly, let I_{tl} denote the total local-mean interference power in the receiver, and let E_{il} denote the local-mean received energy per symbol due to interference signal i. The local-mean SINR is defined to be $E_{sl}/(N_0+I_{tl}T_c)$. For this analysis, a *local-mean outage* is said to occur if the local-mean SINR is less than a specified threshold Z_l, which may be adjusted to account for the fading statistics and any diversity or rake combining. When the local-mean power control is imperfect, $E_{sl} = E_{s0}\epsilon_0$ and $E_{il} = E_{s0}\epsilon_i$, $i = 1, 2, \ldots, K-1$, where ϵ_0 and ϵ_i are lognormally-distributed random variables with the common variance σ_{le}^2. A derivation similar to that leading to (3-112) indicates that if (3-111) is satisfied, then

$$P_{\text{out}}(K) = Q\left\{ -\frac{\ln[(K-1)G^{-1}Z_l \exp(b^2\sigma_{le}^2/2) + Z_l\gamma_0^{-1}]}{b\sigma_{le}} \right\} \tag{3-128}$$

and P_{out} is calculated by using (3-117) and (3-118). The intercell interference factor g can be determined by setting $\sigma_\eta = \sigma_s$ since the fading statistics do not affect the local-mean SINR. For adequate network performance in practical applications, Z_l must be set much higher than the threshold Z in (3-112) because the local-mean SINR changes much more slowly than the instantaneous SINR.

The following example is used to compare the results of evaluating (3-117), (3-118), and (3-128) with the results obtained by Corazza et al. [27], who used a far more elaborate analysis. Consider a cellular network with three sectors, $Z_l = 7$ dB, $\sigma_s = 6$ dB, and $q = 3/8$ due to the voice activity. Table 1 gives $g = 0.558$. A spectral band of bandwidth $W = 1.25$ MHz is occupied by the direct-sequence CDMA signals. The symbol rate is $1/T_s = 8$ kb/s so that the processing gain is $G = 156.5$. The local-mean signal-to-noise ratio before the despreading is -1 dB and $\gamma_0 = 20.94$ dB after the despreading. Figure 35 shows the local-mean outage probability versus the average number of mobiles per cell, 3μ, which is triple the average number of mobiles per sector. The results of [27] for outage probabilities of 10^{-1}, 10^{-2}, and 10^{-3} are indicated by the open circles. The proximity of these points to the curves indicates that the simple equations (3-117), (3-118), and (3-128) closely approximate the local-mean outage probability.

3.10 Bit-Error-Probability Analysis

Uplink capacity is the number of mobiles per cell that can be accommodated over the uplink at a specified information-bit error rate. Assuming a conventional correlation receiver and typical conditions for cellular communications, the subsequent results indicate that when imperfect power control causes the standard deviation of the received power from each mobile to increase beyond 2 dB, the uplink capacity rapidly collapses. The results are consistent with those obtained by several other authors for slow fading (e.g., [22], [27]), but are examined here for fast-fading conditions also. When the instantaneous signal level cannot be tracked, one might consider

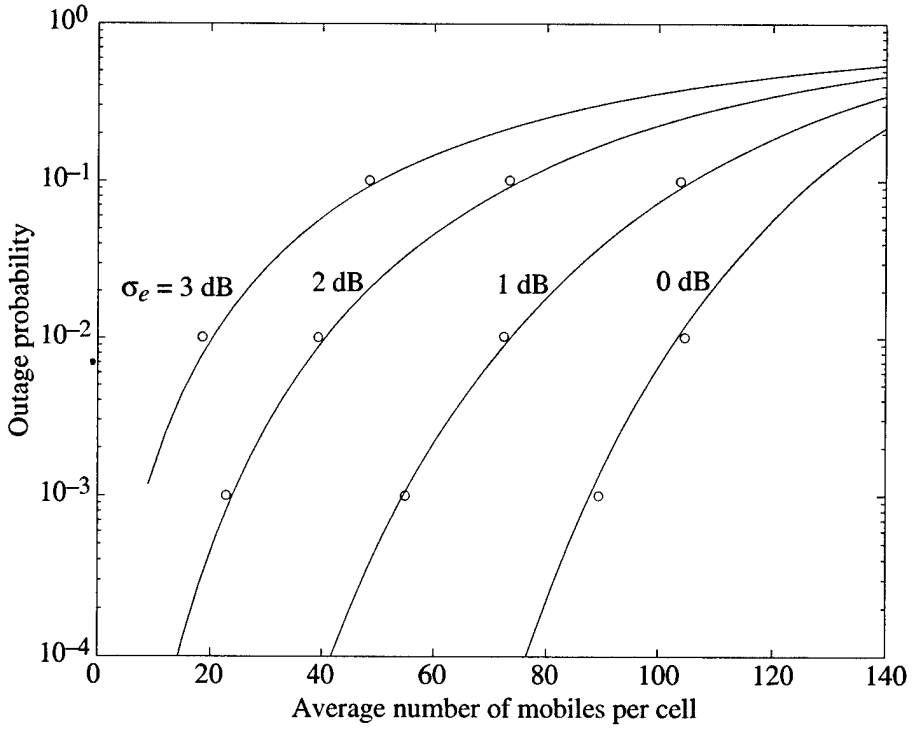


Figure 35. Local-mean outage probability for $Z_l = 7$ dB, $q = 3/8$, $g = 0.558$, $G = 156.5$, and $\gamma_0 = 20.94$ dB with $\sigma_e = 0$, 1, 2, 3 dB. Other theoretical results are indicated by the open circles.

measuring the local-mean power. Accurate local-mean power control eliminates the near-far problem and shadowing effects, but not the effects of the fading. In the subsequent analysis, it is confirmed that tracking the local-mean power is less useful than attempting to track the instantaneous signal level even if the latter results in large errors.

Consider a CDMA cell or sector with K active mobiles. The direct-sequence signals use quadriphase-shift keying (QPSK) modulation. Equation (3-84) indicates that the conditional symbol error probability given E_s and $E_t = I_t T_s$ is approximately given by

$$P_s(E_s, E_t) = Q\left(\sqrt{\frac{2E_s}{N_0 + \sqrt{h}G^{-1}E_t}}\right) \tag{3-129}$$

It is assumed that the distribution of E_s and E_t and the values of G and N_0 are such that (3-83), which is used in the derivation of (3-129), is satisfied with high probability in the subsequent analysis. We consider three models for power control: perfect instantaneous power control (perfect ipc), imperfect instantaneous power control (imperfect ipc) with lognormally distributed errors, and perfect local-mean power control (perfect lmpc).

If the power control is instantaneous and perfect, then $E_i = E_s = E_{s0}$, $i = 1, 2, \ldots, K-1$, and $E_t = (K-1)E_{s0}$. Equation (3-129) implies that the conditional symbol error probability given K is

$$P_s(K) = Q\left(\sqrt{\frac{2}{\gamma_0^{-1} + \sqrt{h}G^{-1}(K-1)}}\right) \quad \text{(perfect ipc)} \tag{3-130}$$

where $\gamma_0 = E_{s0}/N_0$ is the energy-to-noise-density ratio when the power control is perfect. If the power control is imperfect with lognormally distributed errors, then

$$E_s = E_{s0}\epsilon_0 \ , \ E_t = E_{s0}X \tag{3-131}$$

and (3-106) to (3-109) are applicable. If (3-111) is satisfied, then $\bar{X} >> \sigma_x$, and X is well-approximated by $\bar{X}$. Since $\epsilon_0 = \exp(b\xi_0)$ and ξ_0 has a Gaussian density, (3-129) and an integration over this density yield

$$P_s(K) = \int_{-\infty}^{\infty} \frac{\exp(-x^2/2\sigma_e^2)}{\sqrt{2\pi}\sigma_e} Q\left(\sqrt{\frac{2\exp(bx)}{\gamma_0^{-1} + \sqrt{h}G^{-1}(K-1)\exp(b^2\sigma_e^2/2)}}\right) dx$$
(imperfect ipc) (3-132)

Suppose that instead of the instantaneous signal power, the local-mean power averaged over the fast fading is tracked. If this tracking provides perfect power control of the local-mean power at a specific level, then a received signal still exhibits fast fading relative to this level. If the fast fading has a Rayleigh distribution but the fading level is constant over a symbol interval, then the received energy per symbol is $E_s = E_{s0}\epsilon_0$, where ϵ_0 has the exponential probability density function given by (3-119). Therefore, (3-129) implies that the conditional symbol error probability given E_t is

$$\begin{aligned} P_s(E_t) &= \int_0^{\infty} \exp(-x) Q\left(\sqrt{\frac{2x}{\gamma_0^{-1} + \sqrt{h}G^{-1}E_t/E_{s0}}}\right) dx \\ &= \frac{1}{2} - \frac{1}{2}\left(1 + \gamma_0^{-1} + \sqrt{h}G^{-1}E_t/E_{s0}\right)^{-1/2} \end{aligned} \tag{3-133}$$

where the integral is evaluated in the same way as (2-57). The total interference energy E_t is given by (3-131) and (3-106), where each ϵ_i is an independent, exponentially distributed random variable with mean equal to unity. Therefore E_t/E_{s0} has a gamma probability density function given by (A-49) with $N = K - 1$, and for $K \geq 2$ the conditional symbol error probability given K is

$$P_s(K) = \frac{1}{2} - \frac{1}{2}\int_0^{\infty} \frac{x^{K-2}\exp(-x)}{(K-2)!\left(1 + \gamma_0^{-1} + \sqrt{h}G^{-1}x\right)^{1/2}} dx \quad \text{(perfect lmpc)} \tag{3-134}$$

Perfect symbol interleaving is defined as interleaving that causes independent symbol errors in a codeword. Assuming that fast fading enables perfect symbol interleaving, the information-bit error probability for hard-decision decoding can be calculated by substituting (3-130), (3-132), or (3-134) into (2-178), which gives

$$P_{ib}(K) \approx \frac{q}{2(q-1)}\left[\sum_{i=t+1}^{d_m} \frac{d_m}{n}\binom{n}{i} P_s^i (1-P_s)^{n-i} + \sum_{i=d_m+1}^{n} \binom{n-1}{i-1} P_s^i (1-P_s)^{n-i}\right] \tag{3-135}$$

where q is the number of symbols in the alphabet, n is the code length, d_m is the minimum distance, and t is the maximum number of symbol errors that can always be corrected. If r is the code rate and E_b is the energy per bit that is available when the channel symbols are uncoded, then $\gamma_0 = rE_b/N_0$ in (3-130), (3-132), and (3-134). As was done previously, the impact of the intercell interference is modeled by replacing K with $K(1+g)$ in the preceding equations, where g is obtained from Table 1. Averaging over K by using (3-116), we obtain

$$P_{ib} = \sum_{k=1}^{\infty} \frac{\exp(-\lambda)\lambda^k}{[1-\exp(-\lambda)]k!} P_{ib}(k) \tag{3-136}$$

where the equivalent number of mobiles λ is given by (3-118).

Suppose that the fading is slow enough that the interleaving is ineffective and, hence, the error in the instantaneous power control is fixed over a codeword duration. Then an approximation similar to that preceding (3-132) implies that the information-bit error probability for hard-decision decoding of a block code given K is

$$P_{ib}(K) = \int_{-\infty}^{\infty} \frac{\exp(-x^2/2\sigma_e^2)}{\sqrt{2\pi}\sigma_e} P'_{ib}(K, P_s(x))\, dx \tag{3-137}$$

where $P'_{ib}(K, P_s(x))$ is given by (3-135) with P_s replaced by

$$P_s(x) = Q\left(\sqrt{\frac{2\exp(bx)}{\gamma_0^{-1} + \sqrt{h}G^{-1}(K-1)\exp(b^2\sigma_e^2/2)}}\right) \tag{3-138}$$

Equations (3-136) to (3-138) give the information-bit error probability for slow fading.

Plots of the information-bit error probability versus λ for instantaneous power control, γ_0 = 13 dB, G = 128, a rectangular chip waveform with h = 2/3, and various values of σ_e in decibels are illustrated in Figure 36. The block code is the binary BCH (63,30) code, for which d_m = 21 and t = 10. Equations (3-137) and (3-138) are used for slow fading, and (3-130), (3-132), and (3-135) are used for fast fading. When the fading is slow and the interleaving is ineffective, the coding is, as expected, less effective than when the fading is fast and the interleaving is perfect, provided that σ_e remains the same. However, σ_e increases with the fading rate, as shown subsequently. The figure indicates that when $\sigma_e > 2$ dB, there is a severe uplink capacity loss for slow fading and a substantial one for fast fading. The results for other block codes are qualitatively similar.

The use of spatial diversity or, in the presence of frequency-selective fading, a rake receiver will improve the performance of a direct-sequence CDMA system during both slow and fast fading, but the improvement is much greater when the fading is slow. As the fading rate increases, the accuracy of the estimation of the channel parameters used in the rake or diversity combiner becomes more difficult. When the channel-parameter estimation

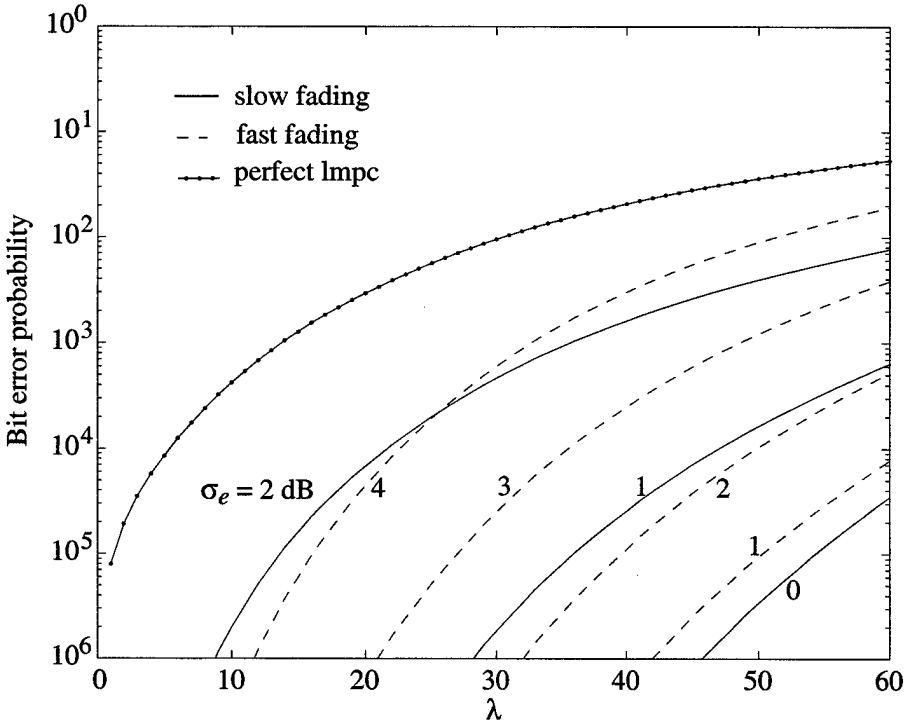


Figure 36. Information-bit error probability for instantaneous power control and perfect local-mean power control, $\gamma_0 = 13$ dB, $G = 128$, and the BCH (63,30) code with various values of σ_e in decibels.

errors are too large to be accommodated, the coherent maximal-ratio combiner must be replaced by the suboptimal noncoherent equal-gain combiner, which does not require the estimation of channel parameters.

In Figure 36, the information-bit error probability is depicted for perfect local-mean power control with the same parameter values and coding as for instantaneous power control. It is assumed that fast fading permits perfect interleaving so that (3-134) and (3-135) are applicable. The figure confirms that tracking of the local-mean power level is an inferior strategy for obtaining a large capacity compared with tracking of the instantaneous power level unless the inaccuracy of the latter is substantial. Another problem with local-mean power control is that it requires time that may be unavailable for sporadic data.

Apart from power control, instantaneous power measurements can be used to facilitate adaptive coding or adaptive transmitter diversity. Both of these techniques require timely information about the impact of the fading, and this information is inherent in the instantaneous power measurements.

3.11 Impact of Doppler Spread on Power-Control Accuracy

When the received instantaneous power of the desired signal from a mobile is tracked, there are four principal error components. They are the quantization error due to the stepping of the transmitted power level, the error introduced in the decoding of the power-control information at the mobile, the error in the power measurement at the base station, and the error caused by the processing and propagation delay. Let $\sigma_q, \sigma_d, \sigma_m$, and σ_p denote the standard deviations of these errors, respectively, expressed in

decibels relative to the received power. Usually, σ_m and σ_p are much larger than σ_q and σ_d [22]. The processing and propagation delay is a source of error because the multipath propagation conditions change during the execution of the closed-loop power-control algorithm.

Assuming that the error sources are independent, the variance of the power-control error can be decomposed as

$$\sigma_e^2 = \sigma_m^2 + \sigma_p^2 + \sigma_q^2 + \sigma_d^2 \tag{3-139}$$

If σ_e is to be less than 2 dB and σ_m is typically more than 1.5 dB, as indicated by the results in [22], then even if σ_q and σ_d are small, $\sigma_p < 1.3$ dB is required. Let v denote the maximum speed of a mobile in the network, f_c the carrier frequency of its direct-sequence transmitted signal, and c the speed of an electromagnetic wave. It is assumed that this signal has a bandwidth that is only a few percent of f_c so that the effect of the bandwidth is negligible. The maximum Doppler shift or Doppler spread is

$$f_d = f_c v/c \tag{3-140}$$

which is proportional to the fading rate. To obtain $\sigma_p < 1.3$ dB requires nearly constant values of the channel attenuation during the processing and propagation delay. Thus, this delay must be much less than the coherence time, which is approximately equal to $1/f_d$, as indicated in (1-40). Examination of attenuation plots for representative multipath scenarios [6] indicates that this delay must be less than α/f_d, where $\alpha \approx 0.1$ or less if $\sigma_p < 1.3$ dB is to be attained. The propagation delay for closed-loop power control is $2d/c$, where d is the distance between the mobile and the base station. Therefore, the processing delay T_p must satisfy

$$T_p < \frac{\alpha}{f_d} - \frac{2d}{c} \tag{3-141}$$

Since T_p must be positive, this inequality and (3-140) imply that $\sigma_p < 1.3$ dB is only possible if $f_c < \alpha c^2/2dv$. Thus, if the carrier frequency or maximum vehicle speed is too high, then the propagation delay alone makes it impossible for the system to attain the required σ_p throughout the network. If $v =$ 25 m/s, $d = 10$ km, $\alpha = 0.1$, and $f_c = 850$ MHz, then (3-141) and (3-140) give $T_p < 1.34$ ms. The IS-95 system, which must accommodate similar parameter values, uses $T_p = 1.25$ ms.

Let p_m denote the measured power level of a received signal in decibels; thus, p_m is an estimate of $10 \log p_0$, where p_0 is the average received signal power from a mobile and the logarithm is to the base 10. Let σ_{m1}^2 denote the variance of an estimate of $\ln p_0$, the natural logarithm of p_0. It follows that the variance of p_m is

$$\sigma_m^2 = (10 \log e)^2 \sigma_{m1}^2 \tag{3-142}$$

It is assumed that power variations in a received signal at the base station are negligible during the measurement interval T_m, which is a large component of the processing delay T_p. Errors in the power measurement occur

because of the presence of multiple-access interference and white Gaussian noise. A lower bound on σ_{m1}^2 can be determined by assuming that the power control is effective enough that the received powers from the mobiles in the cell or sector are approximately equal. The multiple-access interference is modeled as a Gaussian process that increases the one-sided noise power spectral density from N_0 to

$$N_t = N_0 + \frac{p_0}{B}(K-1)(1+g) \tag{3-143}$$

where p_0 is the common signal power of each mobile at the base station and B is the bandwidth of the receiver.

The received signal from a mobile that is to be power-controlled has the form $\sqrt{p_0}s(t)$, where $s(t)$ has unity power. Thus,

$$\int_0^{T_m} s^2(t)dt = T_m \tag{3-144}$$

The received signal can be expressed as

$$\sqrt{p_0}s(t) = \exp\left(\frac{y}{2}\right)s(t) \tag{3-145}$$

where $y = \ln p_0$. The Cramer-Rao bound [28] provides a lower bound on the variance of any unbiased estimate or measurement of $\ln p_0$. This bound and (3-145) give

$$\sigma_{m1}^2 \geq \left\{\frac{2}{N_t}\int_0^{T_m}\left[\frac{\partial}{\partial y}e^{y/2}s(t)\right]^2 dt\right\}^{-1} \tag{3-146}$$

Evaluating (3-146) and using (3-142) and (3-143), we obtain

$$\sigma_m^2 \geq \frac{200(\log e)^2}{p_0 T_m}\left[\frac{N_0}{p_0} + \frac{(K-1)(1+g)}{B}\right] \tag{3-147}$$

Let $T_1 = T_p - T_m$ denote the part of the processing delay in excess of the measurement interval. Substituting (3-139) and (3-141) into (3-147), we obtain

$$\sigma_e^2 > 200(\log e)^2\left(\frac{\alpha}{f_d} - \frac{2d}{c} - T_1\right)^{-1}\left[\frac{N_0}{p_0} + \frac{(K-1)(1+g)}{B}\right] + \sigma_p^2 + \sigma_q^2 + \sigma_d^2 \tag{3-148}$$

This lower bound indicates that σ_e^2 increases with f_d and, hence, the fading rate when the power estimation is ideal.

Inequality (3-148) indicates that an increase in the Doppler spread f_d can be offset by an increase in the bandwidth B. This observation clarifies why third-generation cellular CDMA systems such as WCDMA or cdma 2000 exhibit no more sensitivity to power-control errors than the IS-95 system

despite the substantial increase in the fading rate due to the increased carrier frequency. The physical reason is that an expansion of the bandwidth of the direct-sequence signals allows enough interference suppression to more than compensate for the increased Doppler spread. Furthermore, the potential effect of power-control errors on third-generation CDMA systems is mitigated by the use of convolutional and turbo codes more powerful than the IS-95 codes.

Consider a network of CDMA systems that do not expand the bandwidth when the Doppler spread changes, but adjust T_p so that (3-141) provides a tight bound. Ideal power estimation is assumed so that the lower bound in (3-148) approximates σ_e^2. If the other parameters are unchanged as the Doppler spread changes from f_{d1} to f_{d2}, then σ_e^2 is only affected by the *Doppler factor* defined as

$$D = \frac{f_{d2}}{f_{d1}} \tag{3-149}$$

An example of the impact of the Doppler factor is illustrated in Figure 37, which shows the upper bounds on P_{ib} for instantaneous power control and the BCH (63,30) code. The network experiences slow fading and a Doppler spread $f_{d1} = 100$ Hz. The Doppler factor is $D = 1$. When the Doppler factor is $D = 2$, 3, or 4, perhaps because of increased vehicular speeds, the network is assumed to experience fast fading. The parameter values are $\alpha = 0.1$, $d = 10$ km, $T_1 = 100\ \mu$s, $B = 1/T_c = 1.25$ MHz, $N_0/p_0 = 5\ \mu$s, $\sigma_p^2 + \sigma_q^2 + \sigma_d^2 = 0.5$ (dB)^2, $h = 2/3$, $G = 128$, and $\gamma_0 = p_0 T_s/N_0 = (p_0/N_0)(G/B) = 20 = 13$ dB. The calculations use (3-148), (3-132) and (3-135) to (3-138). In this example, $D \geq 2.5$ causes a significant performance degradation despite the improved time diversity during the fast fading.

When fast fading causes large power-control errors, a direct-sequence CDMA network exhibits a significant performance degradation, notwithstanding the exploitation of time diversity by interleaving and channel coding. Adopting long-term-average instead of instantaneous power control will not cure the problem. A better approach is to increase the bandwidth of the direct-sequence signals. If the bandwidth cannot be increased enough, then the Doppler spread might be reduced by minimizing the carrier frequency of the direct-sequence signals. Another strategy is to limit the size of cells so that the network must cope with the more benign Ricean fading rather than Rayleigh fading, which is more likely to cause large power-control errors.

It follows from (3-147) and $T_p = T_m + T_1$ that a specified σ_m can be attained if

$$T_p \geq \frac{200(\log e)^2}{\sigma_m^2}\left[\left(\frac{N_0}{p_0}\right) + K_1(1+g)\right] + T_1 \tag{3-150}$$

where $K_1 = (K-1)/B$ is the number of interfering active mobiles per unit bandwidth in the cell or sector. Inequalities (3-150) and (3-141) restrict the range of feasible values for T_p.

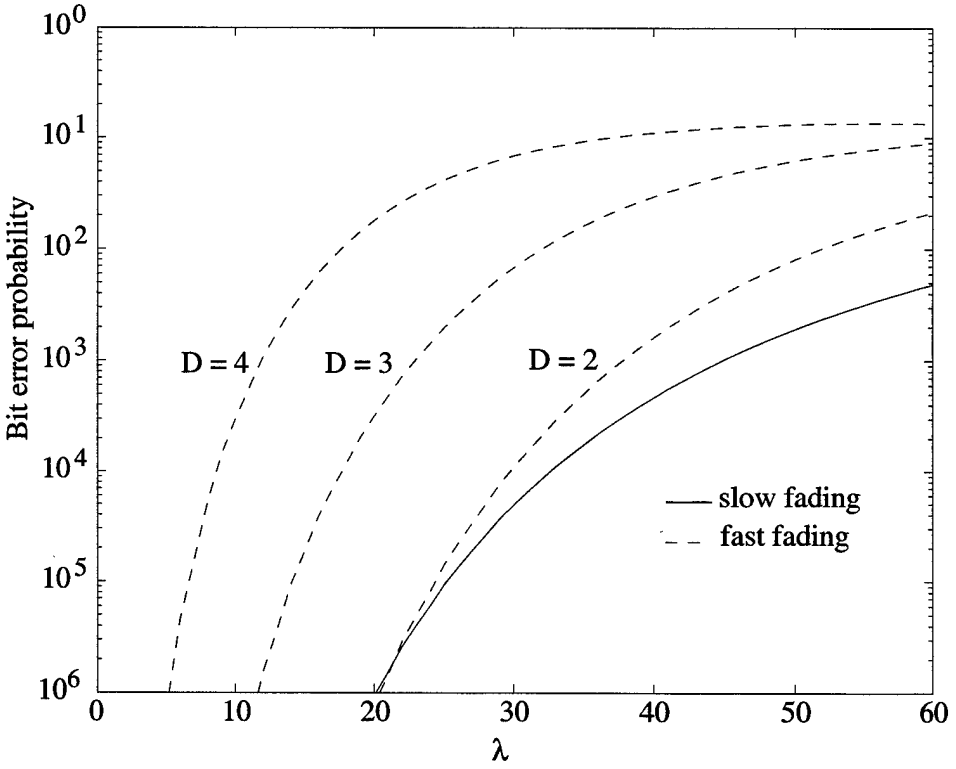


Figure 37. Information-bit error probability for slow fading and fast fading with different Doppler factors D. Instantaneous power control and the BCH (63,30) code are used.

Combining (3-140), (3-141), and (3-150) and assuming that K is large enough that $K_1 \approx K/B$, we conclude that to attain $\sigma_e < 2$ dB for vehicles at speed v or less, an approximate upper bound on the uplink capacity per unit bandwidth in a cell or sector is given by

$$K_1 < (1+g)^{-1}\left[\frac{\sigma_m^2}{200(\log e)^2}\left(\frac{\alpha c}{v f_c} - \frac{2d}{c} - T_1\right) - \left(\frac{N_0}{p_0}\right)\right] \qquad (3\text{-}151)$$

For typical parameter values, this upper bound is approximately inversely proportional to both the carrier frequency f_c and the maximum vehicle speed v.

Figure 38 illustrates the upper bound on the uplink capacity per megahertz as a function of frequency f_c for $\alpha = 0.1$, $\sigma_m = 1.5$ dB, $\sigma_\eta = 8$ dB, $N_0/p_0 = 5$ μs, $T_1 = 100$ μs, and representative values of d and v. Table 1 gives $g = 0.634$. The figure indicates the limitations on K_1 due to power control as the carrier frequency increases if σ_m and the other parameters remain fixed. If K_1 exceeds the upper bound, then the network performance will be severely degraded. The uplink capacity $K_1 B$ can be maintained by expanding the bandwidth.

One might consider circumventing power control by adopting a multiuser detector [29] in the base station in place of the conventional single-user correlation detector. In its optimal form, a multiuser detector can theoretically eliminate the near-far problem along with intracell multiple-access interference, but such a detector is prohibitively complex to implement, especially when long spreading sequences are used. Though suboptimal compared

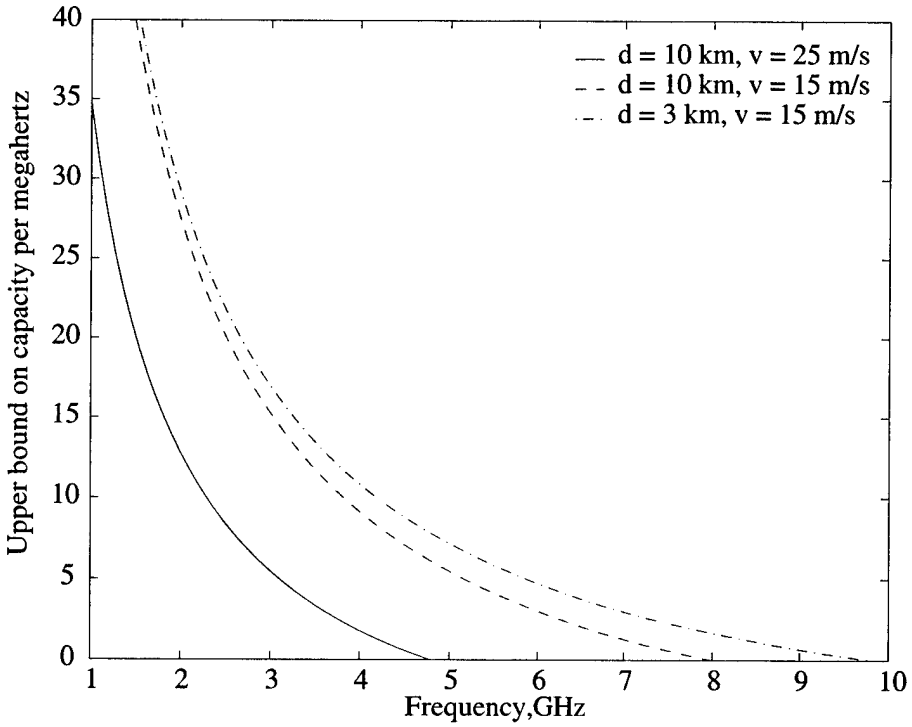


Figure 38. Upper bound on uplink capacity per megahertz for $\alpha = 0.1$, $\sigma_m = 1.5$ dB, $g = 0.634$, $N_0/p_0 = 5$ μs, and $T_1 = 100$ μs.

with ideal multiuser detection, multiuser interference cancellers [30] bear a much more moderate implementation burden and still provide considerable interference suppression. However, it appears that accurate power control is still needed at least for initial synchronization and to avoid overloading the front end of the receiver. Third-generation CDMA systems use adaptive interference cancellation but retain a closed-loop power-control subsystem.

3.12 Downlink Power Control and Outage

Along with all the signals transmitted to mobiles associated with it, a base station transmits a pilot signal over the downlinks. A mobile, which is usually associated with the base station from which it receives the largest pilot signal, uses the pilot to identify a base station or sector, to initiate uplink power control, to estimate the attenuation, phase shift, and delay of each significant multipath component, and to assess the power-allocation requirement of the mobile.

A base station synchronously combines and transmits the pilot and all the signals destined for mobiles associated with it. Consequently, all the signals fade together, and the use of orthogonal spreading sequences will prevent intracell interference and, hence, a near-far problem on a downlink, although there will be interference caused by asynchronously arriving multipath components. The orthogonal sequences can be generated from the rows of a Hadamard matrix defined by (3-17). The orthogonality, the energy-saving sharing of the same pilot at all covered mobiles, and the coherent demodulation of all transmitted signals are major advantages of the

downlinks. However, interference signals from other base stations arrive at a mobile asynchronously and fade independently, thereby significantly degrading performance.

Although there is no near-far problem on the downlinks, power control is still desirable to enhance the received power during severe fading or when a mobile is near a cell edge. However, this power enhancement increases intercell interference. Downlink power control entails power allocation by the base station in a manner that meets the requirements of the individual mobiles associated with it. Let C_{ij} denote the total power received by mobile i from base station j. If this mobile is associated with base station 0, then the SINR at the mobile is

$$\text{SINR} = \frac{\beta \phi_i C_{i0} T_s}{N_0 + \sum_{j=1}^{N_B-1} C_{ij} T_c} \tag{3-152}$$

where N_B is the total number of base stations that produce significant power at mobiles in a cell or sector, β is the fraction of the base-station power that is assigned to mobiles rather than to the pilot, and ϕ_i is the fraction of the total power for mobiles in a cell or sector that is allocated to mobile i. Typically, one might set $\beta = 0.8$, which entails a 1-dB loss due to the pilot. Let R denote the SINR required by network mobiles for acceptable performance. Inverting (3-152), it is found that R is achieved by all mobiles in a cell or sector if

$$\phi_i \geq \frac{R}{\beta C_{i0} G} \left(\frac{N_0}{T_c} + \sum_{j=1}^{N_B-1} C_{ij} \right), \quad i = 1, 2, \ldots, K \tag{3-153}$$

An outage occurs if the demands of all K mobiles in a cell or sector cannot be met simultaneously. Thus, no outage occurs if (3-153) is satisfied and

$$\sum_{i=1}^{K} \chi_i \phi_i \leq 1 \tag{3-154}$$

where χ_i is the voice-activity indicator such that $\chi_i = 1$ with probability q and $\chi_i = 0$ with probability $1 - q$. If the left-hand side of (3-154) is strictly less than unity, then the transmitted power produced by base station 0 can be safely lowered to reduce the interference in other cells or sectors. Combining (3-153) and (3-154), a necessary condition for no outage is

$$\sum_{i=1}^{K} \frac{\chi_i}{C_{i0}} \left(\frac{N_0}{T_c} + \sum_{j=1}^{N_B-1} C_{ij} \right) \leq \frac{\beta G}{R} \tag{3-155}$$

The assignment of mobile i to base station 0 implies the constraint that $C_{ij} \leq C_{i0}$, $j = 1, 2, \ldots, N_B - 1$, except possibly during a soft handoff. A complete performance analysis with this constraint is difficult. Simulation results [26] indicate that the downlink capacity potentially exceeds the uplink capacity if the orthogonal signaling is not undermined by excessive multipath.

3.13 Multicarrier Direct-Sequence CDMA

A *wideband direct-sequence CDMA* signal, which occupies a spectral band that exceeds the coherence bandwidth, can support a large information rate, provides a large processing gain, and can exploit rake combining of multipath signals. An alternative to wideband direct-sequence CDMA is *multicarrier direct-sequence CDMA* in which the available band is partitioned among multiple direct-sequence signals, each of which has a distinct carrier frequency. The main attractions of the multicarrier system are its potential ability to operate over disjoint, noncontiguous spectral regions and its ability to avoid transmissions in spectral regions with strong interference or where the multicarrier signal might interfere with other signals. These features have counterparts in frequency-hopping CDMA systems.

A typical multicarrier system divides a spectral band of bandwidth W into M regions, each of bandwidth W/M approximately equal to the coherence bandwidth [31], [2]. In the transmitter, the product of the data modulation and the spreading waveform, $m(t)p(t)$, simultaneously modulates M carriers, each of which has its frequency in the center of one of the M spectral regions, as illustrated in Figure 39. The receiver comprises M parallel demodulators, one for each carrier, the outputs of which are suitably combined. The total signal power is divided equally among the M carriers. The chip rate and, hence, the processing gain for each carrier of a multicarrier direct-sequence system is reduced by the factor M. However, if strong interference exists in a spectral region, the associated carrier can be omitted and the saved power redistributed among the remaining carriers. Error correcting codes and interleaving can be used to provide both time diversity and coding gain.

Since the M spectral regions are defined so that the fading in each of them is independent and frequency nonselective, rake combining is not possible, but the frequency diversity provided by the M regions can be exploited in a diversity combiner. Whether or not the diversity gain exceeds that of a single-carrier system using the entire spectral band and rake combining depends on the multipath intensity profile.

As an illustration of this issue, we compare the performance of multicarrier and single-carrier direct-sequence systems occupying the same band with Rayleigh fading but no interference. The data modulation is PSK or QPSK. Assuming that the interference among the subcarriers in the receiver is negligible, P_b for the multicarrier system is given by (2-62) and (2-64) with $L = M$ and $\bar{\gamma}/M$ in place of $\bar{\gamma}$ to reflect the power division. Since only white Gaussian noise is present, the processing gain is irrelevant, and the bit error probability P_b for ideal rake combining in the single-carrier receiver is given by (2-148). It is assumed that the largest multipath component in the rake combiner has an average SNR such that $\bar{\gamma}_1 = \bar{\gamma}$ and that $L = 4$ components are present. Figure 40 plots P_b for multicarrier systems with $M = 4$ and 8 and for single-carrier systems with $(\bar{\gamma}_2\ \bar{\gamma}_3\ \bar{\gamma}_4) = (1\ 1\ 0)\bar{\gamma}$ and $(\frac{1}{2}\ \frac{1}{4}\ \frac{1}{8})\bar{\gamma}$. It is observed that the multicarrier system provides a superior diversity gain when M is sufficiently large.

Figure 39. Multicarrier direct-sequence CDMA system: (a) transmitter and (b) receiver.

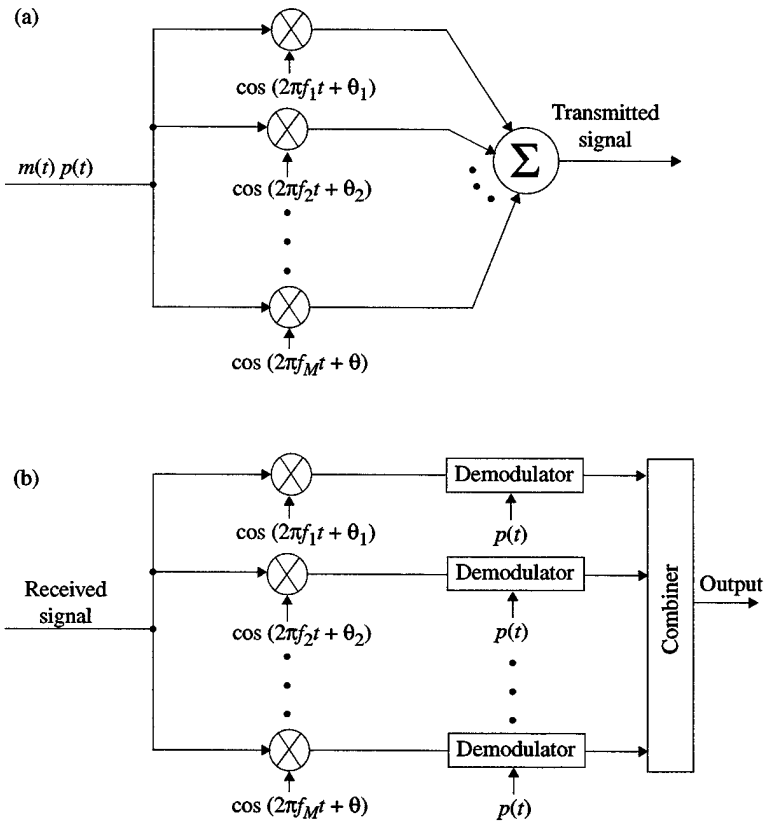
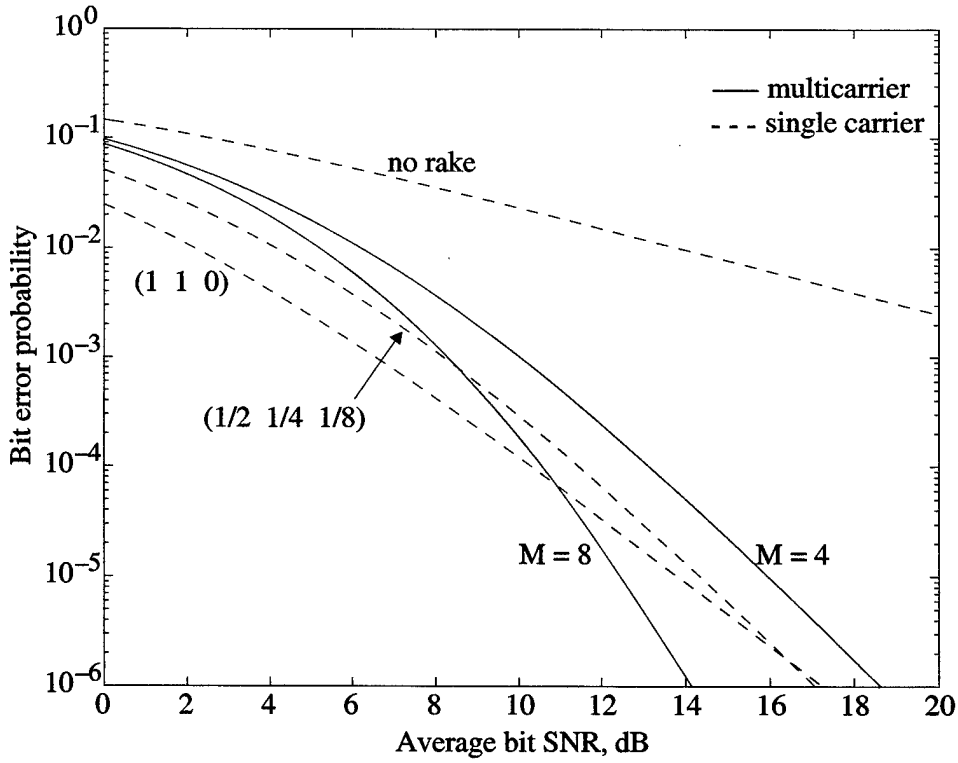


Figure 40. Bit error probability for multicarrier systems with $M = 4$ and 8 and for single-carrier systems with $(\bar{\gamma}_2\bar{\gamma}_3\bar{\gamma}_4) = (1\ 1\ 0)\bar{\gamma}$ and $\left(\frac{1}{2}\ \frac{1}{4}\ \frac{1}{8}\right)\bar{\gamma}$, where $\bar{\gamma}$ is the average bit SNR for both the multicarrier system and the largest multipath component of the single-carrier system.



4. Frequency-Hopping Code-Division Multiple Access

The two principal types of code-division multiple access (CDMA) are direct-sequence CDMA and frequency-hopping CDMA. Two major advantages of frequency hopping are that it can be implemented over a much larger frequency band than it is possible to implement direct-sequence spreading, and that the band can be divided into noncontiguous segments. Another major advantage is that frequency hopping provides resistance to multiple-access interference while not requiring power control to prevent the near-far problem. In direct-sequence systems, accurate power control is crucial but becomes much less effective as the fading rate increases. These advantages of frequency hopping will be decisive in many applications.

Mobile peer-to-peer communications are used in mobile communication networks that possess no supporting infrastructure, fixed or mobile; each user has identical signal processing capability. Peer-to-peer communications have both commercial applications and important military applications, the latter primarily because of their robustness in the presence of node losses. Mobile frequency-hopping CDMA systems [32] are suitable for both peer-to-peer and cellular communication networks. Power control and, hence, current direct-sequence CDMA are not viable for peer-to-peer communications because of the lack of a centralized architecture. Current plans to use multiuser detection in direct-sequence CDMA systems still require power control, which is highly desirable for the synchronization.

Compared with the alternatives, CDMA is advantageous for cellular networks because it eliminates the need for frequency and timeslot coordination among cells, allows complete frequency reuse in all cells, and can fully exploit intermittent voice signals and sectorization. Both frequency-hopping and direct-sequence systems are viable choices for mobile cellular CDMA communications.

Through analysis and simulation, a unified evaluation of the potential performance of both mobile peer-to-peer and sectorized cellular frequency-hopping CDMA systems is presented. The equivalent number of frequency channels and the minimum signal-to-noise ratio (SNR) are defined and shown to be important parameters in understanding and predicting network capacity. The effects of spectral splatter are analyzed. Separated orthogonality is defined and shown to be useful in completely eliminating intracell or intrasector interference. Spatial diversity by postdetection rather than predetection combining is proposed and shown to be invaluable. Noncoherent demodulation by a frequency discriminator rather than parallel matched filters and envelope detectors is proposed and shown to be effective. It is shown that even without exploiting either its natural bandwidth advantage or power control, frequency-hopping CDMA provides an uplink capacity nearly the same as direct-sequence CDMA with realistic power-control imperfections.

4.1 Characteristics and Parameters

The duration of a hop, denoted by T_h, is equal to the sum of the dwell time T_d and the switching time T_{sw}. As illustrated in Figure 41, the *dwell time* is the duration of the frequency-hopping pulse during which the channel symbols are transmitted. The *switching time* is equal to the dead time, when no signal is present, plus the rise and fall times of a pulse. Even if the switching time is absent in the transmitted signal, it will be present in the dehopped signal in the receiver because of the imperfect synchronization of received and receiver-generated waveforms. The nonzero switching time, which may include an intentional guard time, decreases the transmitted symbol duration T_s. If T_{so} is the symbol duration in the absence of frequency hopping, then $T_s = T_{so}(T_d/T_h)$. The reduction in symbol duration expands the transmitted spectrum and thereby reduces the number of frequency channels within a fixed total hopping band. Since the receiver filtering will ensure that rise and fall times of pulses have durations on the order of T_s, the practical assumption that $T_{sw} > T_s$ is made in the subsequent analysis. Thus, each multiple-access interferer transmits in at most one frequency channel during the reception of a single symbol of the desired signal. Since the overhead cost of the nonzero switching time is reduced and equalization symbols can be accommodated, it is preferable to use slow frequency hopping with many symbols per hop, rather than fast frequency hopping, for communications over fading channels. Both slow and fast frequency hopping provide automatic frequency diversity, but slow hopping also requires symbol interleaving over a number of hops.

Even in the absence of fading, there are constraints on the values of the parameters of a frequency-hopping system. It is assumed that the receiver transfer function following the dehopping is approximately rectangular with bandwidth B, which is equal to the bandwidth of a frequency channel. The bandwidth is determined primarily by the percentage of the signal power that must be processed by the demodulator if the demodulated signal distortion and the intersymbol interference are to be negligible. In practice, this percentage must be at least 90 percent and is often more than 95 percent. The relation between B and the symbol duration may be expressed as

$$B = \frac{\zeta}{T_s} \tag{4-1}$$

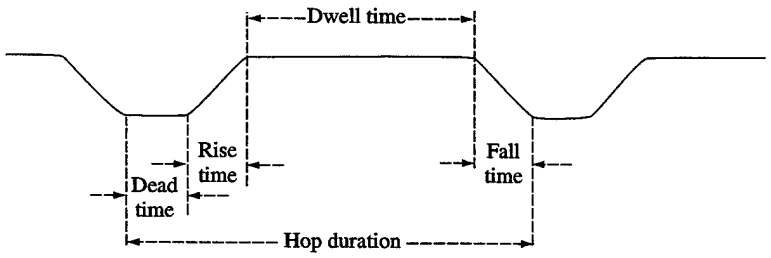


Figure 41. Time durations of a frequency-hopping pulse after the dehopping in the receiver.

where ζ is a constant determined by the signal modulation. For example, if minimum-shift keying is used, then $\zeta = 0.8$ if 90 percent of the signal power is included in a frequency channel, and $\zeta = 1.2$ if 99 percent is included.

Let M denote the number of carriers and frequency channels, W the total bandwidth occupied by the M channels, and F_s the minimum separation between the carriers in a hopset. For full protection against stationary narrowband interference and jamming, it is desirable that $F_s \geq B$ so that the frequency channels are nearly spectrally disjoint.

To obtain the full advantage of block or convolutional error-correcting codes, it is important to interleave the code symbols in a codeword or in a constraint length in such a way that the symbol errors are independent (for hard-decision decoding) or that the symbols are degraded independently (for soft-decision decoding). In frequency-hopping systems operating over a fading channel, the realization of this independence requires certain constraints among the system parameter values. Symbol errors are independent if the fading is independent in each frequency channel and each symbol is transmitted in a different frequency channel. If each of the interleaved code-symbols is transmitted at the same location in each hop dwell time, then adjacent symbols are separated by T_h. Thus, a sufficient condition for nearly independent symbol errors is

$$T_h \geq T_{\text{coh}} \tag{4-2}$$

where T_{coh} is the coherence time of the fading channel. Another sufficient condition for nearly independent symbol errors is

$$F_s \geq B_{\text{coh}} \tag{4-3}$$

where B_{coh} is the coherence bandwidth of the fading channel. For practical mobile communication networks with hopping rates exceeding 100 hops/s, (4-2) is rarely satisfied. In a hopset with a uniform carrier separation, $F_s = W/M \geq B$. Thus, (4-3) implies that the number of frequency channels is constrained by

$$M \leq \frac{W}{\max(B, B_{\text{coh}})} \tag{4-4}$$

if nearly independent symbol errors are to be ensured. If (4-4) is not satisfied, there will be a performance loss due to the correlated symbol errors. If $B < B_{\text{coh}}$, equalization will not be necessary. If $B \geq B_{\text{coh}}$, either equalization may be used or a multicarrier modulation may be combined with the frequency hopping.

Let n denote the length of a block codeword or the constraint length of a convolutional code. Let T_{del} denote the maximum tolerable processing delay. Since the delay caused by coding and ideal interleaving over n hops is $(n-1)T_h + T_s$ and n distinct frequencies are desired,

$$n \leq \min\left(M, 1 + \frac{T_{\text{del}} - T_s}{T_h}\right) \tag{4-5}$$

is required. If (4-5) is not satisfied, then nonideal interleaving is necessary, and some performance degradation results.

Spectral splatter is the interference produced in frequency channels other than the one being used by a frequency-hopping pulse. It is caused by the time-limited nature of transmitted pulses. The degree to which spectral splatter may cause errors depends primarily on F_s, the percentage of the signal power included in a frequency channel, and the spectral characteristics of the signal modulation. One can usually choose these variables so that with high probability only pulses in adjacent channels produce a significant amount of spectral splatter in a frequency channel.

The *adjacent splatter ratio* K_s is the ratio of the power due to spectral splatter from an adjacent channel to the corresponding power that arrives at the receiver in that channel. For example, if B is the bandwidth of a frequency channel that includes 97 percent of the signal power and $F_s \geq B$, then no more than 1.5 percent of the power from a transmitted pulse can enter an adjacent channel on one side of the frequency channel used by the pulse; therefore, $K_s \leq 0.015$. A given maximum value of K_s can be reduced by an increase in F_s, but eventually the value of M must be reduced if W is fixed. As a result, the rate at which users hop into the same channel increases. This increase may cancel any improvement due to the reduction of the spectral splatter. The opposite procedure (reducing F_s and B so that more frequency channels become available) increases not only the spectral splatter but also signal distortion and intersymbol interference, so the amount of useful reduction is limited.

It is assumed that L omnidirectional antennas are deployed to achieve spatial diversity at the mobiles. The antennas are separated from each other by several wavelengths, so that the fading of both the desired signal and the interfering signals at one antenna is independent of the fading at the other antennas. A few wavelengths are adequate because mobiles, in contrast to base stations, tend to receive superpositions of reflected waves arriving from many random angles. Because of practical physical constraints, spatial diversity will ordinarily be effective only if the carrier frequencies exceed roughly 1 GHz. Polarization diversity and other forms of adaptive array processing are alternatives.

One method of combining antenna outputs is predetection combining, which requires the estimation of the signal and interference-plus-noise power levels at each antenna for maximal-ratio combining or selection diversity and requires the cophasing of the L antenna outputs for maximal-ratio or coherent equal-gain combining. Since the relative phases and power levels of the signals at the L antennas change after every hop, it is almost always impractical to implement predetection combining. As a much more practical alternative, a receiver can combine the demodulated outputs rather than the signals from the L antennas. This postdetection combining eliminates the cophasing and does not require the time alignment of L signals in practical applications because any misalignment is much smaller than a symbol duration. The estimation of power levels can be eliminated

by the use of a fixed combining rule, such as equal-gain or square-law combining.

In the receiver of a frequency-hopping system, each antenna output is dehopped and filtered. The interference plus noise in each dehopped signal is approximated by independent bandlimited white Gaussian noise, with equivalent power given by

$$\sigma_1^2 = \sigma_n^2 + \sum_{i=1}^{K} p_{ui} \tag{4-6}$$

where σ_n^2 is the thermal noise power, K is the number of active interfering mobiles, and p_{ui} is the local-mean interference power received from mobile i. The Gaussian model is reasonable, especially for large numbers of mobiles, because the interference signals are asynchronous, fade independently, and experience different Doppler shifts. The total interference power is approximately uniform (white) over the receiver passband following dehopping if $BT_s = \zeta \leq 1$. The L diversity antennas are assumed to be close enough to each other that the power-law losses and shadowing are nearly the same, and thus the local-mean power from a mobile is the same at each antenna. Each active interfering mobile may actually represent a cluster of mobiles. In this cluster, some discipline such as carrier-sense multiple access is used to ensure that there is at most one transmitted signal at any time.

The desired signal is assumed to experience frequency-nonselective Rayleigh fading. The Rayleigh fading model is appropriate under the pessimistic assumption that the propagation paths are often obstructed, and thus, the power of the direct line-of-sight signal is small compared with the reflected signal power. Frequency-nonselective fading occurs if $B < B_{\text{coh}}$. Rayleigh fading may be negligible if mobile speeds are very low, which would occur if each mobile consisted of a person walking. Shadowing would still occur but would be slowly varying over time.

To avoid spectral spreading due to amplifier nonlinearities, it is desirable for the signal modulation to have a constant envelope. Noncoherent demodulation is nearly always a practical necessity in frequency-hopping systems unless the hopping band is narrow. Accordingly, good modulation candidates are differential phase-shift keying (DPSK) and minimum-shift keying (MSK) or some other form of spectrally compact continuous-phase frequency-shift keying (CPFSK) or continuous-phase modulation (CPM). Although the classical theory indicates that DPSK provides an intrinsic performance advantage over noncoherent MSK, this advantage is illusory in a frequency-hopping system because of the large bandwidth requirement of DPSK [3]. If the total hopping bandwidth W is fixed, the number of frequency channels available for DPSK signaling is much smaller than it is for MSK signaling. This reduction in frequency channels largely or entirely offsets the intrinsic performance advantage of DPSK, which is small when postdetection combining is used, as shown subsequently. Alternatively, if the bandwidth B is fixed, a DPSK signal will experience more distortion

and spectral splatter than an MSK signal. Any pulse shaping of the DPSK signals will alter its constant envelope. Another disadvantage of DPSK is due to the usual lack of phase coherence from hop to hop, which necessitates an extra phase-reference symbol at the start of every hop dwell time and thereby causes a performance loss. Thus, DPSK does not appear to be as suitable a means of modulation as MSK, CPFSK, or CPM for frequency-hopping communications.

Spectrally compact CPFSK signals do not have enough frequency shift to be demodulated by classical noncoherent demodulators with parallel matched filters and envelope detectors, but can be demodulated by a frequency discriminator. We consider binary MSK with discriminator demodulation. For postdetection diversity, the outputs of L discriminators are weighted and combined, as shown in Figure 12. The weighting is by the square of the envelope at the input to each discriminator. When the desired signal undergoes independent Rayleigh fading at each antenna and the channel parameters remain constant for at least one symbol duration, a calculation using the results of Adachi and Parsons [33] yields the symbol-error probability

$$P_s = \binom{2L-1}{L} \left(\frac{1}{4} + \frac{1}{3}\zeta^2\right)^L (\bar{\rho})^{-L} \tag{4-7}$$

where $\zeta = BT_s$, $\bar{\rho} = p_s/\sigma_1^2$, and p_s is the local-mean power of the desired signal. A comparison of this equation with (2-67) and (2-101) when $\zeta = 1$ so that $\bar{\rho} = \bar{\gamma}$ verifies that MSK with discriminator demodulation and square-law postdetection combining provides nearly the same performance as DPSK. The slowly varying nature of shadowing ensures that P_s is almost always nearly constant over T_{del} in practical systems. The information-bit error rate following hard-decision decoding can be calculated from P_s with well-known equations. The theoretical loss due to using postdetection rather than predetection combining is less than a decibel [33].

Let d represent the *duty factor*, which is defined as the probability that an interferer using the same frequency will degrade the reception of a symbol. Thus, $d = q_1 q_2$ is the product of the probability q_1 that an interferer is transmitting and the probability q_2 that a significant portion of the interferer's transmitted waveform occurs during the symbol interval. The probability q_2 is upper bounded and well approximated by the probability that there is any overlap in time of the interference and the symbol interval. For synchronized frequency-hopping pulses, $q_2 = 1$. Since $T_{sw} > T_s$, it follows from elementary probability that for unsynchronized frequency-hopping pulses, $q_2 \approx (T_d + T_s)/T_h$. For voice communications with voice-activity detection, $q_1 = 0.4$ is a typical value.

4.2 Peer-to-Peer Simulation Results

This section considers a peer-to-peer network of independent, identical, frequency-hopping systems that have omnidirectional antennas, generate

the same output power, share the same carriers and frequency channels, and are nearly stationary in location over a single symbol duration. Since for peer-to-peer communications it is assumed that an interfering mobile may transmit in any frequency channel with equal probability, the probability that power from an interferer enters the transmission channel of the desired signal is

$$P_t = \frac{d}{M} \tag{4-8}$$

It is assumed that M is sufficiently large that we may neglect the fact that a channel at one of the ends of the hopping band has only one adjacent channel instead of two. Consequently, the probability that the power from an interferer enters one of the two adjacent channels of the desired signal is

$$P_a = \frac{2d}{M} \tag{4-9}$$

The probability that the power enters neither the transmission channel nor the adjacent channels is $(1 - 3d/M)$. These equations make it apparent that the performance of a frequency-hopping system depends primarily on the ratio $M_1 = M/d$. This ratio is called the *equivalent number of channels* because any decrease in the duty factor has the same impact as an increase in the number of frequency channels; what matters most for performance is this ratio.

In the simulation, the locations of the mobiles are assumed to be uniformly distributed in a circular region surrounding a specific mobile receiver, as illustrated in Figure 42. Therefore, the radial distance of a mobile from the receiver has the probability distribution function

$$G(r) = \frac{r^2}{R^2} \quad , \quad 0 < r \leq R \tag{4-10}$$

where R is the radius of the circle. The distance of the desired mobile is randomly selected according to this distribution with $R = R_0$, where R_0 is

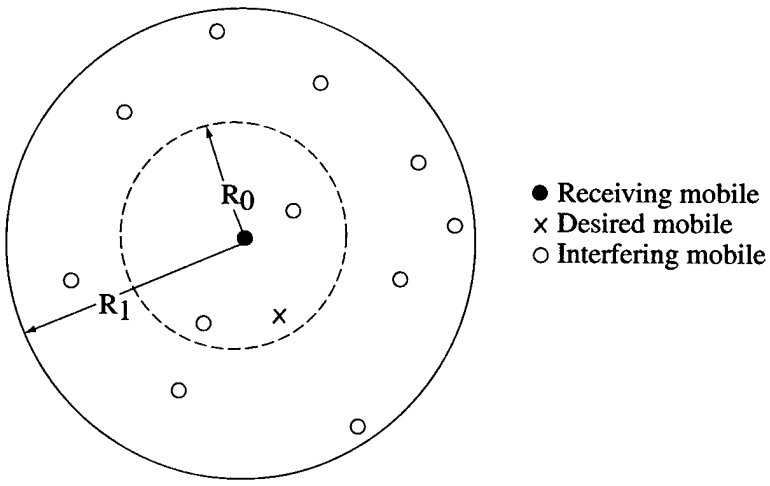


Figure 42. Geometry of a peer-to-peer communication network.

the maximum communication range and corresponds to a received area-mean signal power equal to p_0. The distance of each interfering mobile is randomly selected according to this distribution with $R = R_1$. The selected distance of the desired mobile is substituted into (1-3) of Section 1.1 as the value of r, and then (1-3) is used to randomly select the local-mean power of the desired signal at the receiver. The probabilities given by (4-8) and (4-9) are used to determine if an interfering mobile produces power in the transmission channel or in one of the adjacent channels of the desired signal. If the power enters the transmission channel, then the power level is randomly selected according to (1-3) with the distance of the mobile substituted. If the power enters one of the adjacent channels, then the potential local-mean power level is first randomly selected via (1-3) and then multiplied by K_s to determine the net interference power p_{ui} that appears in (4-6). The effects of p_0 and σ_n^2 are determined solely by the *minimum area-mean SNR*, which occurs at the maximum range $r = R_0$ of the desired signal and is equal to p_0/σ_n^2. Once the local-mean power levels and the noise power are calculated, the symbol error probability P_s is calculated with (4-6) and (4-7). Each simulation experiment was repeated for 10,000 trials, with different randomly selected mobile locations in each trial. The performance measure is the *spatial reliability*, which is defined as the fraction of trials for which P_s is less than a specified performance threshold E. The appropriate value of the threshold depends on the desired information-bit error probability and the error-correcting code. The spatial reliability is essentially the probability that an outage does not occur.

Figures 43 to 45 depict the results of three simulation experiments for peer-to-peer networks. The figures plot the spatial reliability as a function of K for various values of L, assuming (4-7), Rayleigh fading, MSK, and that $\beta = 4$, $\sigma_s = 8$ dB, $E = 0.01$, $\zeta = 1$, $K_s = 0.015$, $R_0 = 1$, and $R_1 = 2$. The value of K_s results from assuming contiguous frequency channels so that $F_s = B$. The units of R_0 and R_1 are immaterial to the calculation of the spatial diversity.

Figure 43 provides a baseline with which the other figures may be compared. For this figure, the assumptions are that $M_1 = 250$, and the minimum area-mean SNR = 20 dB. The number of equivalent frequency channels M_1 could model voice communications with $M = 90$ channels and $d = 0.36$; alternatively, it could model continuous data communications with $M = 225$ and $d = 0.9$. The figure illustrates the dramatic performance improvement provided by dual spatial diversity when Rayleigh fading occurs. Further increases in diversity yield diminishing returns. One can assess the impact of the spectral splatter in this example by setting $K_s = 0$ and observing the change in the spatial reliability. The change is small, and nearly imperceptible if $K < 25$.

Figure 43. Spatial reliability for $M_1 = 250$ and minimum area-mean SNR = 20 dB.

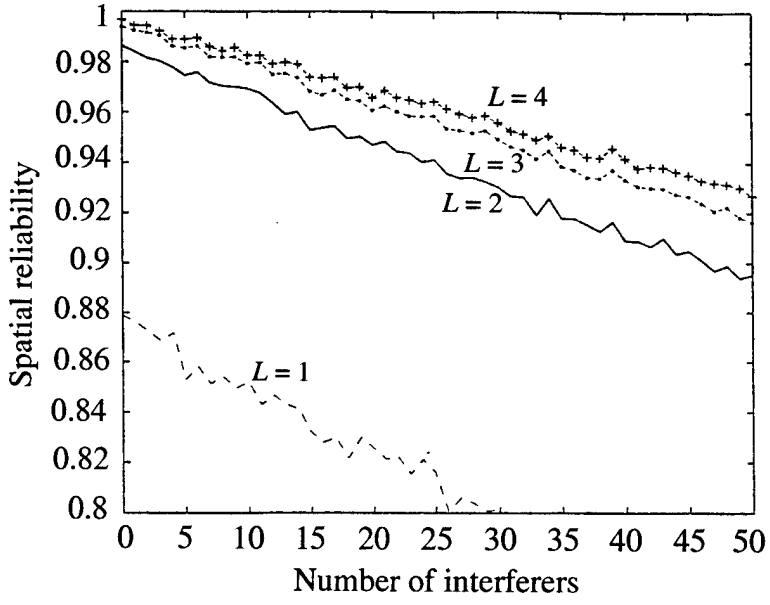


Figure 44 illustrates the effect of increasing the number of equivalent channels to $M_1 = 500$. Let the *capacity* of the network be defined as the maximum number of interfering mobiles for which the spatial reliability exceeds 0.95. Figures 43 and 44 and other simulation results indicate that for the parameter values selected, the capacity C for dual spatial diversity is approximately proportional to M_1; specifically, $C \approx 0.07\, M_1$ for $100 \leq M_1 \leq 1000$. If E is increased to 0.02, the capacity for dual spatial diversity increases by approximately 20 percent.

Figure 45 illustrates the sensitivity of the network to an increase in the minimum area-mean SNR, which may be due to a change in p_0 or σ_n^2. For no spatial diversity or dual diversity, a substantial performance improvement occurs when the minimum area-mean SNR = 25 dB. Other simulation results indicate that a decrease in the minimum area-mean SNR below 20 dB severely degrades performance.

Since (4-7) relates P_s to $\bar{\rho}$, the local-mean signal-to-interference-plus-noise ratio (SINR), the spatial reliability has an alternative and equivalent definition as the fraction of trials for which the SINR exceeds a specified threshold Z_l. Thus, the curves labeled $L = 1$, 2, 3, and 4 in Figures 43 to 45 (and later in Figures 47 to 51) correspond to $Z_l = 17.7$ dB, 10.0 dB, 7.7 dB, and 6.5 dB, respectively.

4.3 Cellular Systems

In a cellular network, each base station assigns separate directional sector antennas or separate outputs of a phased array to cover disjoint angular sectors in both the transmitting and receiving modes. Typically, there are three sectors, and $2\pi/3$ radians are in each angular sector. The mobile antennas are assumed to be omnidirectional. Ideal sector antennas have uniform gain over the covered sector and negligible sidelobes. With these

Figure 44. Spatial reliability for $M_1 = 500$ and minimum area-mean SNR = 20 dB.

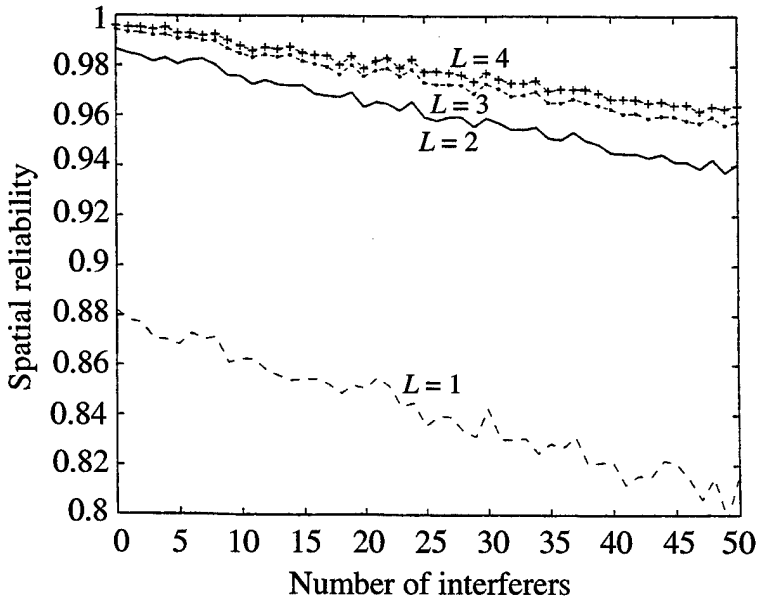
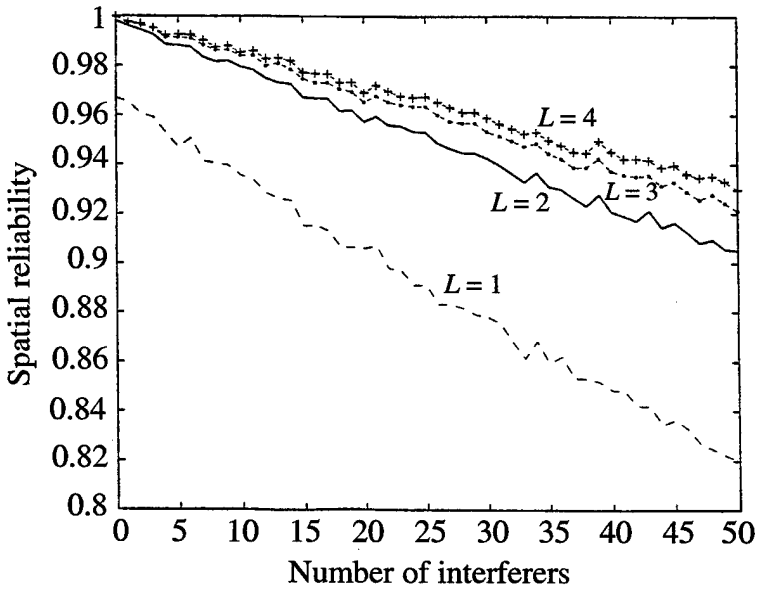


Figure 45. Spatial reliability for $M_1 = 250$ and minimum area-mean SNR = 25 dB.



antennas, only mobiles in the covered sector can cause multiple-access interference on an uplink from a mobile to a base station, and the number of interfering signals on the link is reduced by a factor s approximately equal to the number of sectors. Only the antenna serving a cell sector oriented toward a mobile can cause multiple-access interference on a downlink from the controlling base station to a mobile. Therefore, the number of interfering signals is reduced approximately by a factor s on both the uplinks and downlinks. Practical sector antennas have patterns with sidelobes that extend into adjacent sectors, but the performance degradation due to overlapping sectors is significant only for a small percentage of mobile locations. Ideal sector antennas are assumed in the subsequent simulation.

Spatial diversity may be obtained through the deployment of L antennas in each mobile and L antenna elements for each sector antenna of each base station. The antennas are separated from each other enough that the fading of both the desired signal and the interfering signals at one antenna is independent of the fading at the other antennas. A few wavelengths are adequate for a mobile because it tends to receive superpositions of reflected waves arriving from many random angles. Many wavelengths separation may be necessary for a base station located at a high position, and polarization diversity may sometimes be a more practical means of obtaining diversity.

In a cellular network, the frequency-hopping patterns can be chosen so that at any given instant in time, the frequencies of the mobiles within a cell sector are all different and, hence, the received signals are all orthogonal if the mobile transmissions are properly synchronized. Exact synchronization on a downlink is possible because a common timing is available. The advancing or retarding of the transmit times of the mobiles enables the arrival times at the base station of the uplink signals to be synchronized. The switching time or guard time between frequency-hopping pulses must be large enough to ensure that neither a small synchronization error nor multipath signals can subvert the orthogonality. The appropriate transmit times of a mobile can be determined from position information provided by the Global Positioning System and the known location of the base station. Alternatively, the transmit times can be determined from arrival-time measurements at the base station that are sent to the mobile. These measurements may be based on the adaptive thresholding [16] of the leading and/or trailing edges of a sequence of frequency-hopping pulses.

Let N_s denote the number of mobiles assigned to a cell sector. To ensure orthogonality of N_s received signals within a cell sector, a simple procedure is to generate a periodic frequency-hopping pattern that does not repeat until all the carrier frequencies in a hopset of size $M \geq N_s$ have been used. Mobile n is assigned this pattern with a delay of $n-1$ hop durations, where $n = 1, 2, \ldots, N_s$. If the patterns associated with different sectors are all drawn from a set of one-coincidence sequences [34], then any two signals from different cells or sectors will collide in frequency at a base station at most once during the period of the hopping patterns. However, the use of one-coincidence sequences throughout a network requires frequency planning, which may be too costly in some applications.

It is possible to ensure not only the orthogonality of N_s signals in a sector but also that the received carrier frequencies in any two patterns are separated by at least νB, where ν is a positive integer, so that the spectral splatter is greatly reduced or negligible. Let $k = 0, 1, 2, \ldots, M-1$ label the hopset frequencies in ascending order. Suppose that a frequency-hopping pattern is generated that does not repeat until all the carrier frequencies in a hopset of size $M \geq \nu N_s$ have been used. When mobile 1 hops to frequency k, mobile n hops to frequency $[k + \nu (n-1)]$ modulo M. Frequency-hopping signals that use frequencies determined by this procedure are called *separated*

orthogonal signals. Choosing $\nu = 2$ will generally be adequate because spectral splatter from channels that are not adjacent will be nearly always insignificant if a spectrally compact data modulation is used.

Frequency-hopping CDMA networks largely avoid the near-far problem by continually changing the carrier frequencies so that frequency collisions become brief, unusual events. Thus, power control in a frequency-hopping CDMA network is unnecessary, and all mobiles may transmit at the same power level. When power control is used, it tends to benefit signals from mobiles far from an associated sector antenna while degrading signals from mobiles close to it. Simulation results [35] indicate that even perfect power control typically increases system capacity by only a small amount. There are good reasons to forego this slight potential advantage and not use power control. The required overhead may be excessive. If geolocation of mobiles is done by using measurements at two or more base stations, then the power control may result in significantly less signal power arriving at one or more base stations and the consequent loss of geolocation accuracy.

Consider communications between a base station and a mobile assigned to sector A of a particular cell, as illustrated in Figure 46 for a hexagonal grid of cells. Because of orthogonality, no other signal in sector A will use the same carrier frequency at the same time and thereby cause interference in the transmission channel (current frequency channel) of either the uplink or downlink. Consider another sector covered by the sector antenna of sector A, for example, sector B. Assuming that an interfering signal may independently use any frequency in the network hopset with equal probability, the probability that a mobile in the covered sector produces interference in the transmission channel of the uplink and degrades a particular symbol is

$$P_m = \frac{dN_s}{M} \tag{4-11}$$

This equation also gives the probability that a sector antenna serving another sector that is oriented toward the desired mobile degrades a symbol by producing interference in the transmission channel of the downlink. Because of orthogonality within each sector, no more than one signal from a

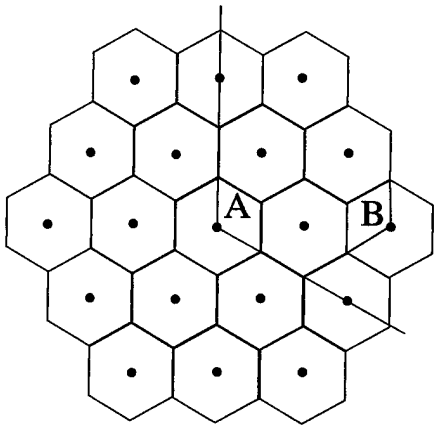


Figure 46. Hexagonal grid of cells. Communicators are in sector A. Sector B is an interfering sector.

sector will produce interference in the transmission channel of either link. A sector with mobiles that may interfere with communications over an uplink or a sector with an antenna that may produce interference over a downlink is called an *interfering sector.*

It is assumed that M is sufficiently large that we may neglect the fact that a channel at one of the ends of the hopping band has only one adjacent channel within the band instead of two. Let $N_1 = 1$ if a signal from an interfering sector uses the transmission channel of communicators in sector A; let $N_1 = 0$ if it does not. The probability that $N_1 = 1$ is N_s/M. The $N_s - N_1$ interference signals from a sector that do not enter the transmission channel are assumed to be randomly distributed among the $M - 1$ frequency channels excluding the transmission channel. There are $\binom{M-1}{N_s-N_1}$ ways to choose the channels with interference signals. There are $\binom{2}{1}$ ways to choose one of the two adjacent channels to have an interference signal and $\binom{M-3}{N_s-N_1-1}$ ways to choose $N_s - N_1 - 1$ channels with interference signals out of the $M - 3$ channels excluding both the transmission channel and the adjacent channels. The probability that an adjacent channel with an interference signal actually receives interference power is q_1. Similarly, there is one way to choose both adjacent channels with interference signals and $\binom{M-3}{N_s-N_1-2}$ ways to choose $N_s - N_1 - 2$ channels with interference signals out of $M - 3$ channels. The probability that exactly one of the two adjacent channels with interference signals actually receives interference power is $2q_1(1 - q_1)$. Because of the sector synchronization, either all of the signals from a sector overlap a desired symbol with probability q_2 or none of them do. Therefore, the probability that a symbol is degraded by interference in exactly one of the adjacent channels of the communicators is

$$\begin{aligned} P_{a1} &= \frac{\binom{2}{1}\binom{M-3}{N_s-N_1-1}}{\binom{M-1}{N_s-N_1}} q_1 q_2 + \frac{\binom{M-3}{N_s-N_1-2}}{\binom{M-1}{N_s-N_1}} 2\, q_1(1-q_1) q_2 \\ &= \frac{2d(N_s - N_1)}{(M-1)(M-2)}[M - 2 - q_1(N_s - N_1 - 1)], \quad M \geq N_s \end{aligned} \tag{4-12}$$

Similarly, the probability that a symbol is degraded by interference in both adjacent channels is

$$\begin{aligned} P_{a2} &= \frac{\binom{M-3}{N_s-N_1-2}}{\binom{M-1}{N_s-N_1}} q_1^2 q_2 \\ &= \frac{dq_1\,(N_s - N_1)\,(N_s - N_1 - 1)}{(M-1)(M-2)}, \quad M \geq N_s \end{aligned} \tag{4-13}$$

For adjacent-channel interference from within sector A, P_{a1} and P_{a2} are given by the same equations with $N_1 = 1$ to reflect the fact that one of the mobiles is the communicating mobile.

Suppose that separated orthogonal frequency-hopping patterns with $\nu = 2$ are used. There is no adjacent-channel interference from sector A. If a signal from an interfering sector B uses the transmission channel so that $N_1 = 1$, an event with probability N_s/M, then the carrier separation of the signals

generated in sector B ensures that there is no adjacent-channel interference from sector B. Suppose that no signal from sector B uses the transmission channel so that $N_1 = 0$. Interference in exactly one adjacent channel results if the transmission channel of the desired signal in sector A, which may be any of $M - N_s$ channels, is located next to one of the two end channels of the set of N_s separated channels being used in sector B, neglecting hopset end effects. It also results if the transmission channel is located between two separated channels, of which only one is currently being used in sector B, again neglecting hopset end effects. Therefore, the probability that a symbol is degraded by interference in exactly one of the adjacent channels of the communicators is

$$\begin{aligned} P_{a1} &= \left[\frac{2q_1}{M-N_s} + \frac{N_s-1}{M-N_s}q_1(1-q_1)\right] q_2 \\ &= \frac{d}{M-N_s}\left[(N_s-1)(1-q_1)+2\right], \quad M \geq 2N_s, \, N_1 = 0 \end{aligned} \tag{4-14}$$

Interference in both adjacent channels results if the transmission channel is located between two separated channels of sector B and both are being used, neglecting hopset end effects. Therefore, the probability that a symbol is degraded by interference in both adjacent channels is

$$P_{a2} = \frac{dq_1(N_s-1)}{M-N_s}, \quad M \geq 2N_s, \, N_1 = 0 \tag{4-15}$$

4.4 Cellular Simulation Results

In the simulation, the spatial configuration consists of a hexagonal grid of cells with base stations at their centers. Each cell has a radius R_0 from its center to a corner. A central cell is surrounded by an inner concentric tier of 6 cells and an outer concentric tier of 12 cells, as depicted in Figure 46. Other tiers are assumed to generate insignificant interference in the central cell. An equal number of mobiles, each transmitting at the same power level, is located in each sector and served by that sector's antenna. This assumption is pessimistic since slightly improved performance may be possible if a mobile is served by the sector antenna providing a signal with the least attenuation and if hysteresis effects during handoffs are not too severe. Each signal transmitted by a sector antenna is allocated the same power. The set of frequency-hopping patterns used in each sector is assumed to be selected independently of the other sectors. Since the parameter R_0 in (1-1) and (1-3) is equal to the maximum communication range, p_0 is the minimum received area-mean power of a desired signal. The location of each mobile within a sector is assumed to be uniformly distributed.

In each simulation trial for communications in sector A of the central cell, the location of the desired mobile is randomly selected according to the uniform distribution. The selected distance of the desired mobile is substituted into (1-3) as the value of r, and then (1-3) is used to randomly select the local-mean power of the desired signal at the receiver. Each transmitting and receiving beam produced by a sector antenna is assumed to have a constant gain over its sector and zero gain elsewhere.

For an uplink of sector A, interference is assumed to arrive from mobiles within sector A, mobiles in the 6 sectors of the two cells in the inner tier that were covered by the beam of sector A, and mobiles in the 11 complete sectors and 2 half-sectors of the five cells in the outer tier completely or partially covered by the beam. The 2 half-sectors are approximated by an additional complete sector in the outer tier. Equations (4-11) to (4-15) are used to determine if a sector contains mobiles that produce power in the transmission channel or in one or both of the adjacent channels. If the sector does, then the locations of the three or fewer interfering mobiles are randomly selected according to the uniform distribution, and their distances from the central cell's base station are computed.

For a downlink of sector A, interference is assumed to arrive from the facing sector antenna of each cell in the two surrounding tiers. Equations (4-11) to (4-15) are used to determine if a signal generated by an interfering sector antenna produces power in the transmission channel or the adjacent channels of the desired signal. If so, then the distance between the sector antenna and the desired mobile is computed.

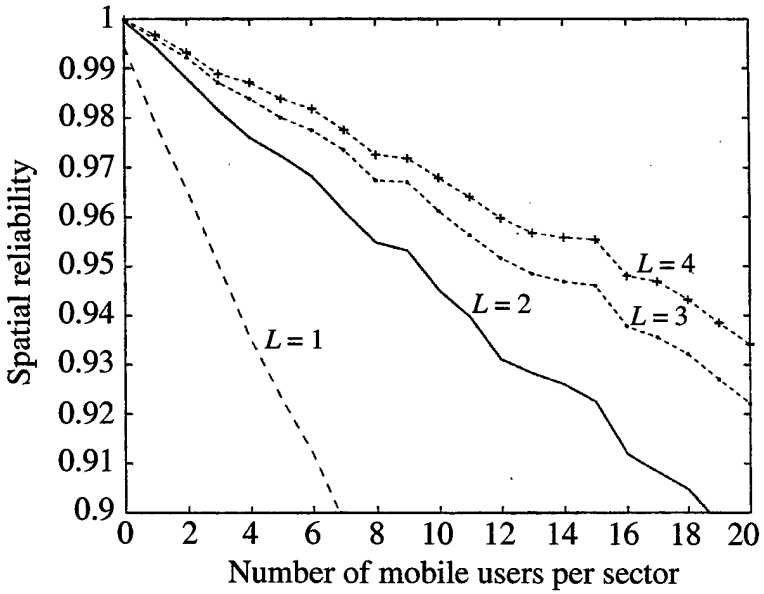
If the power from an interferer enters the transmission channel, then the power level is randomly selected according to (1-3), with the appropriate distance substituted. If the power enters an adjacent channel, then the potential local-mean power level is first randomly selected via (1-3) and then multiplied by K_s to determine the net interference power p_{ui} that appears in (4-6). The shadowing parameter σ_s is assumed to be the same for all signals originating from all cells. The effects of p_0 and σ_n^2 are determined solely by p_0/σ_n^2, the minimum area-mean SNR. Since only ratios affect the performance, the numerical value of R_0 in the simulation is immaterial and is set equal to unity.

Once the local-mean power levels and the noise power are calculated, the symbol error probability is calculated with (4-6) and (4-7). Each simulation experiment was repeated for 20,000 trials, with different randomly selected mobile locations in each trial. The performance measure is the spatial reliability, which is a function of $\bar{\rho}$, the SINR. The appropriate value of the threshold E depends on the desired information-bit error probability and the error-correcting code.

Figures 47 to 50 depict the results of four simulation experiments for the uplinks of a cellular network. The figures plot spatial reliability as a function of N_s for various values of L, assuming MSK, three sectors, and that $\beta = 4$, $q_1 = 0.4$, $q_2 = 1.0$, $\sigma_s = 8$ dB, $E = 0.01$, $\zeta = 1$, and $K_s = 0.015$. The value of K_s results from assuming contiguous frequency channels so that $F_s = B$.

Figure 47 provides a baseline with which other figures may be compared. For this figure, separated orthogonal frequency hopping with $\nu = 2$, $M = 100$, and minimum area-mean SNR = 30 dB are assumed. The figure illustrates the dramatic performance improvement provided by dual spatial diversity when Rayleigh fading occurs. Further increases in diversity yield diminishing returns. One can assess the impact of the spectral splatter in this example by setting $K_s = 0$ and observing the change in spatial

Figure 47. Spatial reliability for uplinks, separated orthogonal hopping, $M = 100$, and minimum area-mean SNR = 30 dB.



reliability. The change is insignificant because by far the most potentially damaging splatter arises from mobiles in the same sector as the desired mobile, and the separated orthogonality has eliminated it.

Figure 48 shows the effect of using orthogonal rather than separated orthogonal frequency hopping. The performance loss is significant in this example and becomes more pronounced as M decreases. When separated orthogonal frequency hopping is used and the spectral splatter is negligible, then the spatial reliability depends primarily on $M_1 = M/d$, the equivalent number of channels. In Figure 47, $M_1 = 250$.

Figure 48. Spatial reliability for uplinks, orthogonal hopping, $M = 100$, and minimum area-mean SNR = 30 dB.

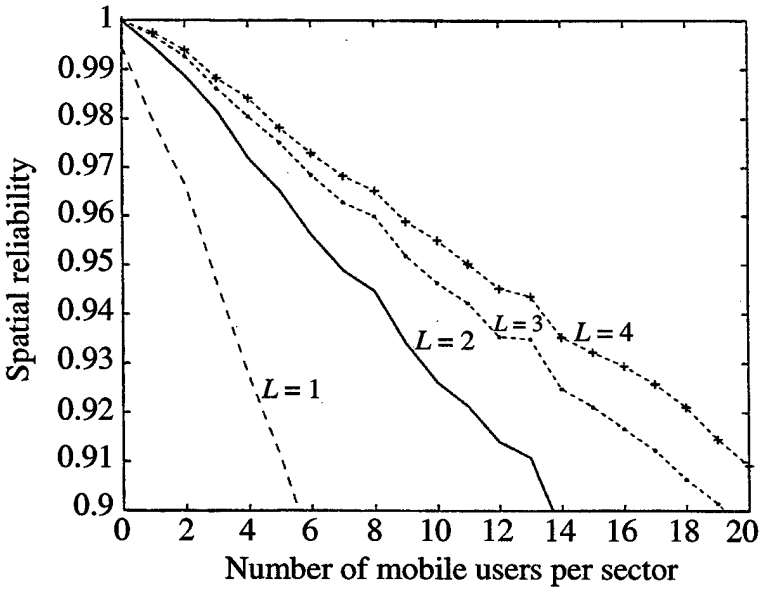


Figure 49 illustrates the effect of increasing M to 200, and hence increasing M_1 to 500. The *uplink capacity* C_u of a cellular network is defined as the maximum number of interfering mobiles per cell for which the spatial reliability exceeds 0.95. Figures 47 and 49 and other simulation results indicate that for three sectors per cell, dual diversity, and the other parameter values selected, the uplink capacity is $C_u \approx 0.108\, M_1$ for $50 \leq M_1 \leq 1000$. This equation is sensitive to parameter variations. If the shadowing standard deviation σ_s is lowered to 6 dB, it is found that C_u increases by roughly 57 percent. Alternatively, if the threshold E is raised to 0.04, corresponding to SINR = 7 dB, it is found that C_u increases by roughly 59 percent.

Figure 50 illustrates the sensitivity of the network to a decrease in the minimum area-mean SNR, which may be due to a change in either p_0 or σ_n^2. A substantial performance loss occurs when the minimum area-mean SNR is reduced to 20 dB, particularly for no spatial diversity or dual diversity. Other simulation results indicate that an increase in the minimum area-mean SNR beyond 30 dB barely improves performance.

The downlinks of a cellular network are considered in Figure 51, where the models and parameter values are otherwise the same as in Figure 47. The performance is worse for the downlinks of Figure 51 than for the uplinks of Figure 47 because of the relative proximity of some of the interfering sector antennas to the desired mobile. The *downlink capacity* C_d, which is defined analogously to the uplink capacity, is $C_d \approx 0.072\, M_1$ for $50 \leq M_1 \leq 1000$. A more realistic comparison of the downlinks and uplinks must take into account the differences between the high-power amplifiers and low-noise amplifiers in the base station and those in the mobiles. Assuming a net 10 dB advantage in the minimum area-mean SNR for the downlinks, Figures 50 and 51 provide a performance comparison of the two links. The performance of the downlinks is still slightly worse if $L \geq 2$ and $N_s \geq 4$.

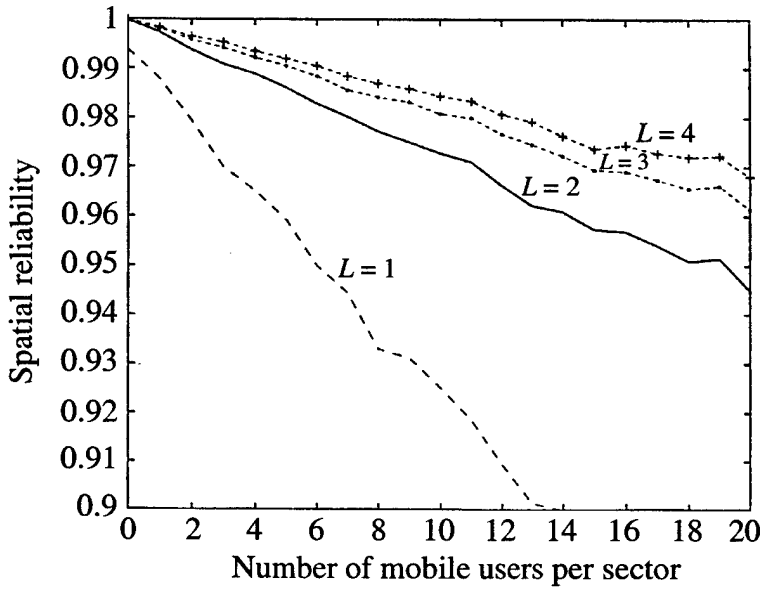


Figure 49. Spatial reliability for uplinks, separated orthogonal hopping, $M = 200$, and minimum area-mean SNR = 30 dB.

Figure 50. Spatial reliability for uplinks, separated orthogonal hopping, $M = 100$, and minimum area-mean SNR = 20 dB.

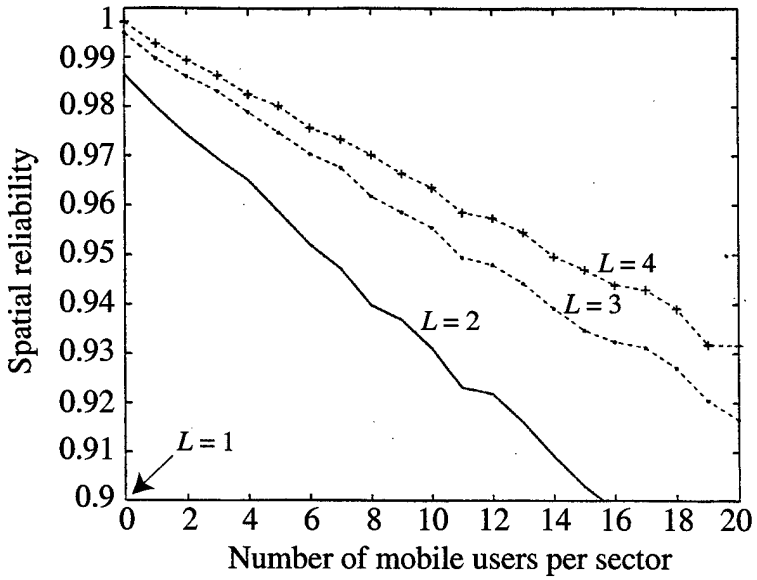
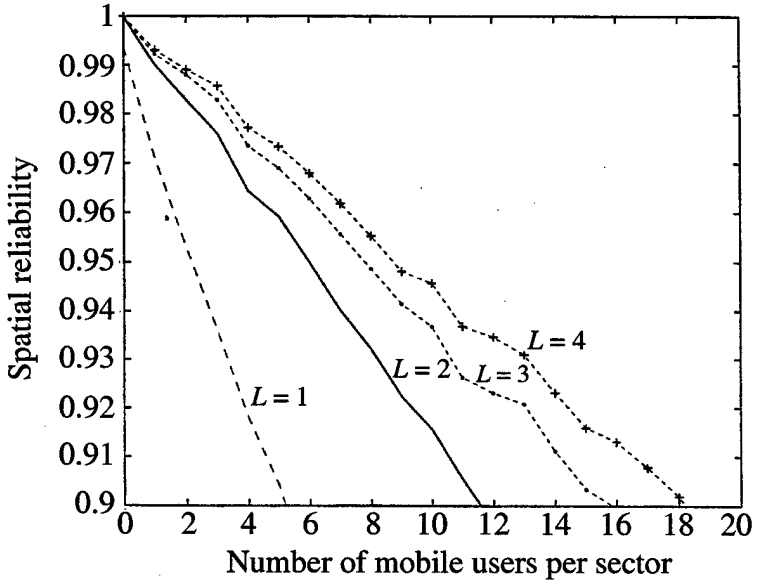


Figure 51. Spatial reliability for downlinks, separated orthogonal hopping, $M = 100$, and minimum area-mean SNR = 30 dB.



Compared with direct-sequence systems, frequency-hopping systems have a bandwidth advantage in that frequency hopping over a large, possibly noncontiguous, spectral band is as practical as direct-sequence spreading over a much smaller, necessarily contiguous, spectral band. Even deprived of its bandwidth advantage, as well as power control and the use of one-coincidence codes, frequency-hopping CDMA can provide nearly the same multiple-access capacity over the uplinks as direct-sequence CDMA subject to realistic power-control imperfections. For a numerical example, consider a cellular network with three sectors, shadowing standard deviation $\sigma_s = 6$ dB, and $d = q_1 = 3/8$ due to the voice activity. A contiguous spectral band of bandwidth $W = 1.25$ MHz is occupied by the CDMA signals. The symbol rate is $1/T_s = 8$ kb/s so that the processing gain is

156.5 for direct-sequence CDMA, and the number of frequency channels for frequency-hopping CDMA with $\zeta = BT_s = 1$ is $M = 156$. For direct-sequence CDMA, it is assumed that p_0 is the received power at the base station from all associated mobiles when the power control is perfect and that the SNR before the despreading is -1 dB. Therefore, the SNR is 20.94 dB after the despreading. For frequency-hopping CDMA without power control, the minimum area-mean SNR is assumed to be 20.94 dB. The uplink capacity C_u is calculated as the number of mobiles per cell that can be accommodated while maintaining an SINR above a specified threshold Z with 95 percent probability. For frequency-hopping CDMA with dual diversity and $Z = 10$ dB, it is found that $C_u \approx 60$. For direct-sequence CDMA with dual diversity and coherent phase-shift keying, a comparison of (4-7) with (2-67) and (2-62) indicates that a comparable performance can be obtained when the SINR is roughly 3 dB less. Thus, the threshold for direct-sequence CDMA is set at $Z = 7$ dB. Using (3-112) with $Z = 7$ dB, it is found that $C_u \approx 60$ when the power-control error has $\sigma_e = 2$ dB.

For coherent demodulation of a signal that hops over a wide band to be a practical possibility in a fading environment, either a pilot signal must be available or the dwell time must be large enough that a small portion of it can be dedicated to carrier synchronization. In the latter case, the degradation due to the dedicated portion and the occasional failure to achieve carrier synchronization for a frequency-hopping pulse must be less than the potential gain due to the coherent demodulation, which is large. If ideal coherent demodulation is assumed in the preceding example so that $Z = 7$ dB, then it is found that $C_u \approx 108$, an increase of 80 percent. This uplink capacity is approximately obtained by direct-sequence CDMA with $Z = 7$ dB when $\sigma_e = 0.4$ dB, an impractically low value.

4.5 Summary

The performance of frequency-hopping CDMA communications in a mobile peer-to-peer network is greatly improved by the use of spatial diversity, which usually requires carrier frequencies in excess of 1 GHz. A crucial parameter is the number of equivalent frequency channels, which can be increased not only by an increase in the number of frequency channels, but also by a decrease in the duty factor of the network users. The data modulation method that is most suitable appears to be MSK or some other form of CPFSK or CPM. For these modulations, $BT_s \approx 1$, and the scenario modeled, the spectral splatter from adjacent channels, is not an important factor if the number of interferers is much smaller than the number of equivalent channels.

For a specified sectorization, diversity, and waveform, the capacity of a cellular frequency-hopping CDMA network is approximately proportional to the equivalent number of frequency channels. Thus, a desired capacity can be attained by choosing a sufficiently large number of frequency channels. A major advantage of frequency hopping is that these channels do not have to be spectrally contiguous but can be scattered throughout a large spectral band. Another advantage is that power control is not required. Its absence

allows a substantial reduction of system complexity and overhead cost and facilitates geolocation. Sectorization, orthogonality, and dual diversity are invaluable, but higher levels of diversity offer sharply decreasing gains. If spectral splatter is a problem, separated orthogonal signaling can be used to eliminate it. The overall limit on the capacity of a frequency-hopping CDMA network appears to be set more by the downlinks than the uplinks.

Frequency hopping may be added to almost any communication system to strengthen it against interference or fading. For example, the set of carriers used in a multicarrier direct-sequence CDMA system or the subcarriers of an *orthogonal frequency-division-multiplexing* (OFDM) system may be hopped. The purpose of OFDM is to enable high data-rate communications without an elaborate equalization system by transmitting symbols simultaneously over a number of subcarriers.

Appendix A. Probability Distributions

A-1 Chi-Square Distribution

Consider the random variable

$$Z = \sum_{i=1}^{N} A_i^2 \tag{A-1}$$

where the $\{A_i\}$ are independent Gaussian random variables with means $\{m_i\}$ and common variance σ^2. The random variable Z is said to have a *noncentral chi-square (χ^2) distribution* with N degrees of freedom and a *noncentral parameter*

$$\lambda = \sum_{i=1}^{N} m_i^2 \tag{A-2}$$

To derive the probability density function of Z, we first note that each A_i has the density function

$$f_{A_i}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m_i)^2}{2\sigma^2}\right] \tag{A-3}$$

From elementary probability, the density of $Y_i = A_i^2$ is

$$f_{Y_i}(x) = \frac{1}{2\sqrt{x}}[f_{A_i}(\sqrt{x}) + f_{A_i}(-\sqrt{x})]\, u(x) \tag{A-4}$$

where $u(x) = 1, x \geq 0$, and $u(x) = 0, x < 0$. Substituting (A-3) into (A-4), expanding the exponentials, and simplifying, we obtain the density

$$f_{Y_i}(x) = \frac{1}{\sqrt{2\pi x}\sigma} \exp\left(-\frac{x+m_i^2}{2\sigma^2}\right) \cosh\left(\frac{m_i\sqrt{x}}{\sigma^2}\right) u(x) \tag{A-5}$$

The characteristic function of a random variable X is defined as

$$C_X(j\nu) = E[e^{j\nu X}] = \int_{-\infty}^{\infty} f_X(x) \exp(j\nu x)\, dx \tag{A-6}$$

where $j = \sqrt{-1}$, and $f_X(x)$ is the density of X. Since $C_X(j\nu)$ is the conjugate Fourier transform of $f_X(x)$,

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_X(j\nu) \exp(-j\nu x)\, d\nu \tag{A-7}$$

From Laplace or Fourier transform tables, it is found that the characteristic function of $f_{Y_i}(x)$ is

$$C_{Y_i}(j\nu) = \frac{\exp[jm_i^2\nu/(1-j2\sigma^2\nu)]}{(1-j^2\sigma^2\nu)^{1/2}} \tag{A-8}$$

The characteristic function of a sum of independent random variables is equal to the product of the individual characteristic functions. Because Z is the sum of the Y_i, the characteristic function of Z is

$$C_Z(j\nu) = \frac{\exp[j\lambda\nu/(1-j2\sigma^2\nu)]}{(1-j^2\sigma^2\nu)^{N/2}} \tag{A-9}$$

where we have used (A-2). From (A-9), (A-7), and Laplace or Fourier transform tables, we obtain the probability density function of *noncentral* χ^2 *random variable* with N degrees of freedom and a noncentral parameter λ:

$$f_Z(x) = \frac{1}{2\sigma^2}\left(\frac{x}{\lambda}\right)^{(N-2)/4} \exp\left[-\frac{x+\lambda}{2\sigma^2}\right] I_{N/2-1}\left(\frac{\sqrt{x\lambda}}{\sigma^2}\right) u(x) \tag{A-10}$$

where $I_\alpha(\ )$ is the modified Bessel function of the first kind and order α. This function may be represented by

$$I_\alpha(x) = \sum_{i=0}^{\infty} \frac{(x/2)^{\alpha+2i}}{i!\,\Gamma(\alpha+i+1)} \tag{A-11}$$

where the gamma function is defined as

$$\Gamma(x) = \int_0^\infty y^{x-1}\exp(-y)dy \ , \ x > 0 \tag{A-12}$$

The probability distribution function of a noncentral χ^2 random variable is

$$F_Z(x) = \int_0^x \frac{1}{2\sigma^2}\left(\frac{y}{\lambda}\right)^{(N-2)/4} \exp\left(-\frac{y+\lambda}{2\sigma^2}\right) I_{N/2-1}\left(\frac{\sqrt{y\lambda}}{\sigma^2}\right) dy \ , \ x \geq 0 \tag{A-13}$$

When N is even so that $N/2$ is an integer, then using $F_Z(\infty) = 1$ and a change of variables in (A-13) yields

$$F_Z(x) = 1 - Q_{N/2}\left(\frac{\sqrt{\lambda}}{\sigma}, \frac{\sqrt{x}}{\sigma}\right) \ , \ x \geq 0 \tag{A-14}$$

where the *generalized Q-function* is defined as

$$Q_m(\alpha,\beta) = \int_\beta^\infty x\left(\frac{x}{\alpha}\right)^{m-1} \exp\left(-\frac{x^2+\alpha^2}{2}\right) I_{m-1}(\alpha x)\,dx \tag{A-15}$$

and m is an integer. Since $Q_m(\alpha,0) = 1$, it follows that $1 - Q_m(\alpha,\beta)$ is an integral with finite limits that can be numerically integrated. However, the

numerical computation of the generalized Q-function is simplified if it is expressed in alternative forms [2].

The mean, variance, and moments of Z can be easily obtained by using (A-1) and the properties of independent Gaussian random variables. The mean and variance of Z are

$$E[Z] = N\sigma^2 + \lambda \tag{A-16}$$

$$\sigma_z^2 = 2N\sigma^4 + 4\lambda\sigma^2 \tag{A-17}$$

where σ^2 is the common variance of the $\{A_i\}$.

From (A-9), it follows that the sum of two independent noncentral χ^2 random variables with N_1 and N_2 degrees of freedom, noncentral parameters λ_1 and λ_2, respectively, and the same parameter σ^2 is a noncentral χ^2 random variable with $N_1 + N_2$ degrees of freedom and noncentral parameter $\lambda_1 + \lambda_2$.

A-2 Central Chi-Square Distribution

To determine the probability density function of Z when the $\{A_i\}$ have zero means, we substitute (A-11) into (A-10) and then take the limit as $\lambda \to 0$. We obtain

$$f_Z(x) = \frac{1}{(2\sigma^2)^{N/2}\Gamma(N/2)} x^{N/2-1} \exp\left(-\frac{x}{2\sigma^2}\right) u(x) \tag{A-18}$$

Alternatively, this equation results if we substitute $\lambda = 0$ into the characteristic function (A-9) and then use (A-7). Equation (A-18) is the probability density function of a *central* χ^2 *random variable* with N degrees of freedom. The probability distribution function is

$$F_Z(x) = \int_0^x \frac{1}{(2\sigma^2)^{N/2}\Gamma(N/2)} y^{N/2-1} \exp\left(-\frac{y}{2\sigma^2}\right) dy, \quad x \geq 0 \tag{A-19}$$

If N is even so that $N/2$ is an integer, then integrating this equation by parts $N/2 - 1$ times yields

$$F_Z(x) = 1 - \exp\left(-\frac{x}{2\sigma^2}\right) \sum_{i=0}^{N/2-1} \frac{1}{i!} \left(\frac{x}{2\sigma^2}\right)^i, \quad x \geq 0 \tag{A-20}$$

By direct integration using (A-18) and (A-12) or from (A-16) and (A-17), the mean and variance of Z are

$$E[Z] = N\sigma^2 \tag{A-21}$$

$$\sigma_z^2 = 2N\sigma^4 \tag{A-22}$$

A-3 Rice Distribution

Consider the random variable

$$R = \sqrt{A_1^2 + A_2^2} \tag{A-23}$$

where A_1 and A_2 are independent Gaussian random variables with means m_1 and m_2, respectively, and a common variance σ^2. The probability distribution function of R must satisfy $F_R(r) = F_Z(r^2)$, where $Z = A_1^2 + A_2^2$ is a χ^2 random variable with two degrees of freedom. Therefore, (A-14) with $N = 2$ implies that

$$F_R(r) = 1 - Q_1\left(\frac{\sqrt{\lambda}}{\sigma}, \frac{r}{\sigma}\right), \quad r \geq 0 \tag{A-24}$$

where $\lambda = m_1^2 + m_2^2$. This function is called the *Rice probability distribution function*. The *Rice probability density function*, which may be obtained by differentiation of (A-24), is

$$f_R(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2 + \lambda}{2\sigma^2}\right) I_0\left(\frac{r\sqrt{\lambda}}{\sigma^2}\right) u(r) \tag{A-25}$$

The moments of even order can be derived from (A-23) and the moments of the independent Gaussian random variables. The second moment is

$$E[R^2] = 2\sigma^2 + \lambda \tag{A-26}$$

In general, moments of the Rice distribution are given by an integration over the density in (A-25). Substituting (A-11) into the integrand, interchanging the summation and integration, changing the integration variable, and using (A-12), we obtain a series that is recognized as a special case of the confluent hypergeometric function. Thus,

$$E[R^n] = (2\sigma^2)^{n/2} \exp\left(-\frac{\lambda}{2\sigma^2}\right) \Gamma\left(1 + \frac{n}{2}\right) {}_1F_1\left(1 + \frac{n}{2}, 1; \frac{\lambda}{2\sigma^2}\right), \quad n \geq 0 \tag{A-27}$$

where the *confluent hypergeometric function* is defined as

$${}_1F_1(\alpha, \beta; x) = \sum_{i=0}^{\infty} \frac{\Gamma(\alpha + i)\Gamma(\beta)x^i}{\Gamma(\alpha)\Gamma(\beta + i)i!}, \quad \beta \neq 0, -1, -2, \ldots \tag{A-28}$$

The Rice density function often arises in the context of a transformation of variables. Let A_1 and A_2 represent independent Gaussian random variables with common variance σ^2 and means m and zero, respectively. Let R and Θ be implicitly defined by $A_1 = R\cos\Theta$ and $A_2 = R\sin\Theta$. Then (A-23) and $\Theta = \tan^{-1}(A_2/A_2)$ describes a transformation of variables. A straightforward calculation yields the joint density function of R and Θ:

$$f_{R,\Theta}(r, \theta) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2 - 2rm\cos\theta + m^2}{2\sigma^2}\right), \quad r \geq 0, \quad |\theta| \leq \pi \tag{A-29}$$

The density function of the envelope R is obtained by integration over θ. Since

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(x \cos u)\, du \tag{A-30}$$

this density function reduces to the Rice density function of (A-25). The density function of the angle Θ is obtained by integrating (A-29) over r. Completing the square in the argument of (A-29), changing variables, and defining

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy = \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) \tag{A-31}$$

where erfc( ) is the complementary error function, we obtain

$$f_\Theta(\theta) = \frac{1}{2\pi} \exp\left(-\frac{m^2}{2\sigma^2}\right) + \frac{m \cos\theta}{\sqrt{2\pi}\sigma} \exp\left(-\frac{m^2 \sin^2\theta}{2\sigma^2}\right) \left[1 - Q\left(\frac{m\cos\theta}{\sigma}\right)\right], \quad |\theta| \le \pi \tag{A-32}$$

Since (A-29) cannot be written as the product of (A-25) and (A-32), the random variables R and Θ are not independent.

Since the density function of (A-25) must integrate to unity, we find that

$$\int_0^{\infty} r \exp\left(-\frac{r^2}{2b^2}\right) I_0\left(\frac{r\sqrt{\lambda}}{b^2}\right) dr = b^2 \exp\left(\frac{\lambda}{2b^2}\right) \tag{A-33}$$

where λ is a positive constant. This equation is useful in calculations involving the Rice density function.

A-4 Rayleigh Distribution

A Rayleigh-distributed random variable is defined by (A-23) when A_1 and A_2 are independent Gaussian random variables with zero means and a common variance σ^2. Since $F_R(r) = F_Z(r^2)$, where Z is a central χ^2 random variable with two degrees of freedom, (A-20) with $N = 2$ implies that the *Rayleigh probability distribution function* is

$$F_R(r) = 1 - \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \ge 0 \tag{A-34}$$

The *Rayleigh probability density function*, which may be obtained by differentiation of (A-34), is

$$f_R(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) u(r) \tag{A-35}$$

By a change of variables in the defining integral, any moment of R can be expressed in terms of the gamma function defined in (A-12). Therefore,

$$E[R^n] = (2\sigma^2)^{n/2} \Gamma\left(1 + \frac{n}{2}\right) \tag{A-36}$$

Certain properties of the gamma function are needed to simplify (A-36). An integration by parts of (A-12) indicates that $\Gamma(1+x) = x\Gamma(x)$. A direct integration yields $\Gamma(1) = 1$. Therefore, when n is an integer, $\Gamma(n) = (n-1)!$. Changing the integration variable by substituting $y = z^2$ in (A-12), it is found that $\Gamma(1/2) = \sqrt{\pi}$.

Using these properties of the gamma function, we obtain the mean and the variance of a Rayleigh-distributed random variable:

$$E[R] = \sqrt{\frac{\pi}{2}}\sigma \tag{A-37}$$

$$\sigma_R^2 = \left(2 - \frac{\pi}{2}\right)\sigma^2 \tag{A-38}$$

Since A_1 and A_2 have zero means, the joint probability density function of the random variables $R = \sqrt{A_1^2 + A_2^2}$ and $\Theta = \tan^{-1}(A_2/A_1)$ is given by (A-29) with $m = 0$. Therefore,

$$f_{R,\Theta}(r,\theta) = \frac{r}{2\pi\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0, \quad |\theta| \leq \pi \tag{A-39}$$

Integration over θ yields (A-35), and integration over r yields the uniform probability density function:

$$f_\Theta(\theta) = \frac{1}{2\pi}, \quad |\theta| \leq \pi \tag{A-40}$$

Since (A-39) equals the product of (A-35) and (A-40), the random variables R and Θ are independent. In terms of these random variables, $A_1 = R\cos\Theta$ and $A_2 = R\sin\Theta$. A straightforward calculation using the independence and densities of R and Θ verifies that A_1 and A_2 are independent, zero-mean, Gaussian random variables with common variance σ^2.

Since the square of a Rayleigh-distributed random variable may be expressed as $R^2 = A_1^2 + A_2^2$, where A_1 and A_2 are zero-mean independent Gaussian random variables with common variance σ^2, R^2 has the distribution of a central chi-square random variable with 2 degrees of freedom. Therefore, (A-18) with $N = 2$ indicates that the square of a Rayleigh-distributed random variable has an exponential probability density function with mean $2\sigma^2$.

A-5 Sum of Independent, Exponentially Distributed Random Variables

Consider the random variable

$$Z = \sum_{i=1}^{N} Y_i \tag{A-41}$$

where the $\{Y_i\}$ are independent, exponentially distributed random variables with unequal positive means $\{m_i\}$. The exponential probability density function of Y_i is

$$f_{Y_i}(x) = \frac{1}{m_i}\exp\left(-\frac{x}{m_i}\right)u(x) \tag{A-42}$$

A straightforward calculation yields the characteristic function

$$C_{Y_i}(j\nu) = \frac{1}{1 - j\nu m_i} \tag{A-43}$$

Since Z is the sum of independent random variables, (A-43) implies that its characteristic function is

$$C_Z(j\nu) = \prod_{i=1}^{N} \frac{1}{1 - j\nu m_i} \tag{A-44}$$

To derive the probability density function of Z, (A-7) is applied after first expanding the right-hand side of (A-44) in a partial-fraction expansion. The result is

$$f_Z(x) = \sum_{i=1}^{N} \frac{B_i}{m_i} \exp\left(-\frac{x}{m_i}\right) u(x) \tag{A-45}$$

where

$$B_i = \begin{cases} \prod\limits_{\substack{k=1 \\ k \neq i}}^{N} \dfrac{m_i}{m_i - m_k} & , \ N \geq 2 \\ 1 & , \ N = 1 \end{cases} \tag{A-46}$$

and $m_i \neq m_k$, $i \neq k$. A direct integration and algebra yields the probability distribution function

$$F_Z(x) = 1 - \sum_{i=1}^{N} B_i \exp\left(-\frac{x}{m_i}\right), \quad x \geq 0 \tag{A-47}$$

Equations (A-45) and (A-12) give

$$E[Z^n] = \Gamma(1+n) \sum_{i=1}^{N} B_i m_i^n, \quad n \geq 0 \tag{A-48}$$

When the $\{m_i\}$ are equal so that $m_i = m$, $1 \leq i \leq N$, then $C_Z(j\nu) = (1 - j\nu m)^{-N}$. Therefore, the probability density function of Z is

$$f_Z(x) = \frac{1}{(N-1)!m^N} x^{N-1} \exp\left(-\frac{x}{m}\right) u(x) \tag{A-49}$$

which is a special case of the *gamma density function*. Successive integration by parts yields

$$F_Z(x) = 1 - \exp\left(-\frac{x}{m}\right) \sum_{i=0}^{N-1} \frac{1}{i!} \left(\frac{x}{m}\right)^i \tag{A-50}$$

From (A-49) and (A-12), the mean and variance of Z are found to be

$$E[Z] = Nm \tag{A-51}$$

$$\sigma_Z^2 = Nm^2 \tag{A-52}$$

Appendix B. Signal Representations

B-1 Hilbert Transform

Consider a real-valued function $g(t)$ defined in the time interval $-\infty < t < \infty$. The *Hilbert transform* of $g(t)$ is defined by

$$H[g(t)] = \hat{g}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(u)}{t-u}\, du \tag{B-1}$$

Because this integrand has a singularity, we define the integral as its Cauchy principal value. Thus,

$$\int_{-\infty}^{\infty} \frac{g(u)}{t-u}\, du = \lim_{\epsilon \to 0} \left[\int_{-\infty}^{t-\epsilon} \frac{g(u)}{t-u}\, du + \int_{t+\epsilon}^{\infty} \frac{g(u)}{t-u}\, du \right] \tag{B-2}$$

provided that the limit exists. Subsequently, integrals are to be interpreted as Cauchy principal values if they contain singularities.

The definition of the Hilbert transform indicates that $\hat{g}(t)$ may be interpreted as the convolution of $g(t)$ with $1/\pi t$. Therefore, $\hat{g}(t)$ results from passing $g(t)$ through a linear filter with an impulse response equal to $1/\pi t$. The transfer function of the filter is given by the Fourier transform

$$\mathcal{F}\left\{\frac{1}{\pi t}\right\} = \int_{-\infty}^{\infty} \frac{\exp(-j2\pi f t)}{\pi t}\, dt \tag{B-3}$$

where $j = \sqrt{-1}$. This integral can be rigorously evaluated by using contour integration. Alternatively, we observe that since $1/t$ is an odd function,

$$\begin{aligned} \mathcal{F}\left\{\frac{1}{\pi t}\right\} &= -2j \int_{0}^{\infty} \frac{\sin 2\pi f t}{\pi t}\, dt \\ &= -j \operatorname{sgn}(f) \end{aligned} \tag{B-4}$$

where $\operatorname{sgn}(f)$ is the *signum function* defined by

$$\operatorname{sgn}(f) = \begin{cases} 1, & f > 0 \\ 0, & f = 0 \\ -1, & f < 0 \end{cases} \tag{B-5}$$

Let $G(f) = \mathcal{F}\{g(t)\}$, and let $\hat{G}(f) = \mathcal{F}\{\hat{g}(t)\}$. Equations (B-1) and (B-4) and the convolution theorem imply that

$$\hat{G}(f) = -j \operatorname{sgn}(f) G(f) \tag{B-6}$$

Because $H[\hat{g}(t)]$ results from passing $g(t)$ through two successive filters, each with transfer function $-j\,\mathrm{sgn}(f)$,

$$H[\hat{g}(t)] = -g(t) \tag{B-7}$$

provided that $G(0) = 0$.

Equation (B-6) indicates that taking the Hilbert transform corresponds to introducing a phase sift of -90 degrees for all positive frequencies and $+90$ degrees for all negative frequencies. Consequently,

$$H[\cos 2\pi f_0 t] = \sin 2\pi f_0 t \tag{B-8}$$

$$H[\sin 2\pi f_0 t] = -\cos 2\pi f_0 t \tag{B-9}$$

These relations can be formally verified by taking the Fourier transform of the left-hand side of (B-8) or (B-9), applying (B-6), and then taking the inverse Fourier transform of the result. If $G(f) = 0$ for $|f| > W$ and $f_0 > W$, the same method yields

$$H[g(t)\cos 2\pi f_0 t] = g(t)\sin 2\pi f_0 t \tag{B-10}$$

$$H[g(t)\sin 2\pi f_0 t] = -g(t)\cos 2\pi f_0 t \tag{B-11}$$

B-2 Analytic Signal and Complex Envelope

A *bandpass signal* is one with a Fourier transform that is negligible except for $f_c - W/2 \leq |f| \leq f_c + W/2$, where $0 \leq W < 2f_c$ and f_c is the center frequency. If $W << f_c$, the bandpass signal is often called a *narrowband signal*. A complex-valued signal with a Fourier transform that is nonzero only for $f > 0$ is called an *analytic signal*.

Consider a bandpass signal $g(t)$ with Fourier transform $G(f)$. The analytic signal $g_a(t)$ associated with $g(t)$ is defined to be the signal with Fourier transform

$$G_a(f) = [1 + \mathrm{sgn}(f)]G(f) \tag{B-12}$$

which is zero for $f \leq 0$ and is confined to the band $|f - f_c| \leq W/2$ when $f > 0$. The inverse Fourier transform of (B-12) and (B-6) imply that

$$g_a(t) = g(t) + j\hat{g}(t) \tag{B-13}$$

The *complex envelope* of $g(t)$ is defined by

$$g_l(t) = g_a(t)\exp[-j2\pi f_c t] \tag{B-14}$$

where f_c is the center frequency if $g(t)$ is a bandpass signal. Since the Fourier transform of $g_l(t)$ is $G_a(f + f_c)$, which occupies the band $|f| \leq W/2$, the complex envelope is a baseband signal that may be regarded as an *equivalent lowpass representation* of $g(t)$. Equations (B-13) and (B-14) imply that $g(t)$ may be expressed in terms of its complex envelope as

$$g(t) = \mathrm{Re}[g_l(t)\exp(j2\pi f_c t)] \tag{B-15}$$

The complex envelope can be decomposed as

$$g_l(t) = g_c(t) + jg_s(t) \tag{B-16}$$

where $g_c(t)$ and $g_s(t)$ are real-valued functions. Therefore, (B-15) yields

$$g(t) = g_c(t)\cos(2\pi f_c t) - g_s(t)\sin(2\pi f_c t) \tag{B-17}$$

Since the two sinusoidal carriers are in phase quadrature, $g_c(t)$ and $g_s(t)$ are called the *in-phase* and *quadrature* components of $g(t)$, respectively. These components are lowpass signals confined to $|f| \leq W/2$.

From Parseval's relation in Fourier analysis and (B-6), we obtain

$$\int_{-\infty}^{\infty} \hat{g}^2(t)\,dt = \int_{-\infty}^{\infty} |\hat{G}(f)|^2\,df = \int_{-\infty}^{\infty} |G(f)|^2\,df = \int_{-\infty}^{\infty} g^2(t)\,dt \tag{B-18}$$

Therefore,

$$\begin{aligned}\int_{-\infty}^{\infty} |g_l(t)|^2\,dt &= \int_{-\infty}^{\infty} |g_a(t)|^2\,dt = \int_{-\infty}^{\infty} g^2(t)\,dt + \int_{-\infty}^{\infty} \hat{g}^2(t)\,dt \\ &= 2\int_{-\infty}^{\infty} g^2(t)\,dt = 2\mathcal{E}\end{aligned} \tag{B-19}$$

where $\mathcal{E}$ denotes the energy of the bandpass signal $g(t)$.

B-3 Direct-Conversion Receiver

Receivers often extract the complex envelope of the desired signal before applying it to a matched filter. The main components in a *direct-conversion* receiver are shown in Figure B-1(a). The spectra of the received signal $g(t)$, the input to the baseband filter, and the complex envelope $g_l(t)$ are depicted in Figure B-1(b). Let $2f(t)$ denote the impulse response of the filter. The output of the filter is

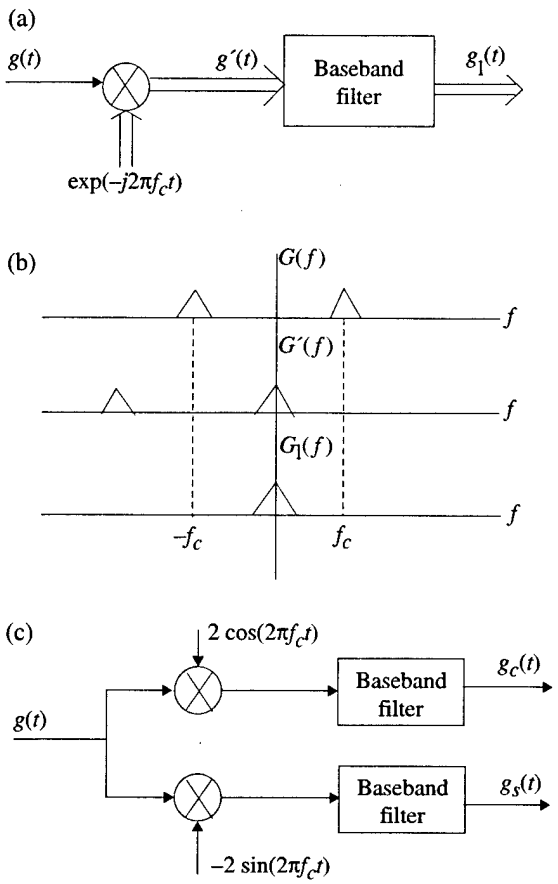
$$y(t) = \int_{-\infty}^{\infty} 2\,g(\tau)\exp(-j2\pi f_c\tau)f(t-\tau)\,d\tau \tag{B-20}$$

Using (B-15) and the fact that $\mathrm{Re}(x) = (x + x^*)/2$, where x^* denotes the complex conjugate of x, we obtain

$$y(t) = \int_{-\infty}^{\infty} g_l(\tau)f(t-\tau)\,d\tau + \int_{-\infty}^{\infty} g_l(\tau)f(t-\tau)\exp(-j4\pi f_c\tau)\,d\tau \tag{B-21}$$

The second term is the Fourier transform of $g_l(\tau)f(t-\tau)$ evaluated at frequency $-2f_c$. Assuming that $g_l(\tau)$ and $f(t-\tau)$ have transforms confined to $|f| < f_c$, their product has a transform confined to $|f| < 2f_c$, and the second term in (B-21) vanishes. If the Fourier transform of $f(t)$ is a constant over the passband of $g_l(t)$, then (B-21) implies that $y(t)$ is proportional to $g_l(t)$, as desired. Figure B-1(c) shows the direct-conversion receiver for real-valued signals.

Figure B-1.
Envelope extraction: (a) direct-conversion receiver, (b) associated spectra, and (c) implementation with real-valued signals.



The direct-conversion receiver alters the character of the noise $n(t)$ entering it. Suppose that $n(t)$ is a zero-mean, white Gaussian noise process with autocorrelation

$$R_n(\tau) = E[n(t)n(t+\tau)] = \frac{N_0}{2}\delta(\tau) \tag{B-22}$$

where $E[\ ]$ denotes the expected value, $\delta(\tau)$ denotes the Dirac delta function, and $N_0/2$ is the two-sided noise-power spectral density. The complex-valued noise at the output of Figure B-1(a) is

$$z(t) = \int_{-\infty}^{\infty} 2n(u)e^{-j2\pi f_c u} f(t-u)\,du \tag{B-23}$$

Since it is a linear function of $n(t)$, $z(t)$ is zero-mean and its real and imaginary parts are jointly Gaussian. The autocorrelation of a wide-sense stationary, complex-valued process $z(t)$ is defined as

$$R_z(\tau) = \frac{1}{2}E[z^*(t)z(t+\tau)] \tag{B-24}$$

Substituting (B-23), interchanging the expectation and integration operations, using (B-22) to evaluate one of the integrals, and then changing variables, we obtain

$$R_z(\tau) = N_0 \int_{-\infty}^{\infty} f(u)f^*(u+\tau)\,du \tag{B-25}$$

If the filter is an ideal bandpass filter with Fourier transform

$$H(f) = \begin{cases} 1, & |f| \leq W \\ 0, & \text{otherwise} \end{cases} \tag{B-26}$$

where $W < f_c$, then the convolution in (B-25) is easily evaluated:

$$R_z(\tau) = N_0 \frac{\sin 2\pi W\tau}{\pi\tau} \tag{B-27}$$

For $\tau > 1/W$, $R_z(\tau)$ has a small value, and the integral of $R_z(\tau)$ over $(-\infty, \infty)$ equals N_0. Thus, as $W \to \infty$, the autocorrelation may be approximated by

$$R_z(\tau) = N_0\delta(\tau) \tag{B-28}$$

This approximation permits major analytical simplifications.

Equations (B-23) and (B-22) imply that

$$E[z(t)z(t+\tau)] = 2N_0 e^{-j4\pi f_c t} \int_{-\infty}^{\infty} e^{j4\pi f_c u} f(u+\tau) f(u)\, du \tag{B-29}$$

Reasoning similar to that following (B-21) lead to

$$E[z(t)z(t+\tau)] = 0 \tag{B-30}$$

A complex-valued stochastic process $z(t)$ that satisfies (B-30) is called a *circularly symmetric* process. Setting $\tau = 0$ in (B-30) and (B-27) yields

$$E[(z^R(t))^2] = E[(z^I(t))^2] = 2N_0 W \tag{B-31}$$

$$E[(z^R(t)z^I(t))] = 0 \tag{B-32}$$

where $z^R(t)$ and $z^I(t)$ are the real and imaginary parts of $z(t)$, respectively.

Appendix C. Jensen's Inequality

A function $g(x)$ defined on an open interval I is *convex* if

$$g(px + (1-p)y) \leq pg(x) + (1-p)g(y) \tag{C-1}$$

for x, y in I and $0 \leq p \leq 1$. Suppose that $g(x)$ has a continuous, nondecreasing derivative $g'(x)$ on I. The inequality is valid if $p = 0$ or 1. If $x \geq y$ and $0 \leq p < 1$,

$$\begin{aligned} g(px + (1-p)y) - g(y) &= \int_y^{px+(1-p)y} g'(z)dz \leq p(x-y)g'(px + (1-p)y) \\ &\leq \frac{p}{1-p}\int_{px+(1-p)y}^{x} g'(z)dz \\ &= \frac{p}{1-p}[g(x) - g(px + (1-p)y)] \end{aligned}$$

Simplifying this result, we obtain (C-1). If $y \geq x$, a similar analysis again yields (C-1). Thus, if $g(x)$ has a continuous, nondecreasing derivative on I, it is convex.

Lemma. If $g(x)$ is a convex function on the open interval I, then

$$g(y) \geq g(x) + g^-(x)(y-x) \tag{C-2}$$

for all y, x in I, where $g^-(x)$ is the left derivative of $g(x)$.

Proof. If $y - x \geq z > 0$, then substituting $p = 1 - z/(y-x)$ into (C-1) gives

$$g(x+z) \leq \left(1 - \frac{z}{y-x}\right) g(x) + \frac{z}{y-x} g(y)$$

which yields

$$\frac{g(x+z) - g(x)}{z} \leq \frac{g(y) - g(x)}{y-x}, \quad y - x \geq z > 0 \tag{C-3}$$

If $v > 0$ and $z > 0$, then (C-1) implies that

$$g(x) \leq \frac{z}{v+z} g(x-v) + \frac{v}{v+z} g(x+z)$$

which yields

$$\frac{g(x) - g(x-v)}{v} \leq \frac{g(x+z) - g(x)}{z}, \quad v, z > 0 \tag{C-4}$$

Inequality (C-3) indicates that the ratio $[g(y) - g(x)]/(y-x)$ decreases monotonically as $y \to x$ from above and (C-4) implies that this ratio has a lower

bound. Therefore, the right derivative $g^+(x)$ exists on I. If $x - y \geq v > 0$, then (C-1) with $p = 1 - v/(x - y)$ implies that

$$g(x - v) \leq \left(1 - \frac{v}{x - y}\right) g(x) + \frac{v}{x - y} g(y)$$

which yields

$$\frac{g(x) - g(y)}{x - y} \leq \frac{g(x) - g(x - v)}{v}, \quad x - y \geq v > 0 \tag{C-5}$$

This inequality indicates that the ratio $[g(x) - g(y)]/(x - y)$ increases monotonically as $y \to x$ from below and (C-4) implies that this ratio has an upper bound. Therefore, the left derivative $g^-(x)$ exists on I, and (C-4) yields

$$g^-(x) \leq g^+(x) \tag{C-6}$$

Taking the limits as $z \to 0$ and $v \to 0$ in (C-3) and (C-5), respectively, and then using (C-6), we find that (C-2) is valid for all y, x in I.□

Jensen's inequality. If X is a random variable with a finite expected value $E[X]$, and $g(\,)$ is a convex function on an open interval containing the range of X, then

$$E[g(X)] \geq g(E[X])$$

Proof. Set $y = X$ and $x = E[X]$ in (C-2), which gives $g(X) \geq g(E[X]) + g^-(E[X])(X - E[X])$. Taking the expected values of the random variables on both sides of this inequality gives Jensen's inequality.□

References

1. G. Stuber, *Principles of Mobile Communication, 2nd ed.* Boston: Kluwer Academic, 2001.

2. M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels.* New York: Wiley, 2000.

3. J. G. Proakis, *Digital Communications, 4th ed.* New York: McGraw-Hill, 2001.

4. S. J. Leon, *Linear Algebra with Applications, 6th ed.* Upper Saddle River, NJ: Prentice-Hall, 2002.

5. J. D. Parsons and J. G. Gardiner, *Mobile Communication Systems.* New York: Halsted Press, 1989.

6. E. A. Lee and D. G. Messerschmitt, *Digital Communication, 2nd ed.* Boston: Kluwer Academic, 1994.

7. J. S. Lee and L. E. Miller, *CDMA Systems Engineering Handbook.* Boston: Artech House, 1998.

8. I. S. Gradsteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products, 6th ed.* San Diego: Academic Press, 2000.

9. M. E. Frerking, *Digital Signal Processing in Communication Systems.* New York: Van Nostrand Reinhold, 1994.

10. S. G. Wilson, *Digital Modulation and Coding.* Upper Saddle River, NJ: Prentice-Hall, 1996.

11. K. Higuchi et al., "Experimental Evaluation of Combined Effect of Coherent Rake Combining and SIR-Based Fast Transmit Power Control for Reverse Link of DS-CDMA Mobile Radio," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1526–1535, August 2000.

12. D. Torrieri, "Simple Formula for Error Probability of Rake Demodulator for Noncoherent Binary Orthogonal Signals and Rayleigh Fading," *IEEE Trans. Commun.*, vol. 50, November 2002.

13. S. Tanaka et al., "Experiments on Coherent Adaptive Antenna Array Diversity for Wideband DS-CDMA Mobile Radio," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1495–1504, August 2000.

14. D. Torrieri, "Information-Bit, Information-Symbol, and Decoded-Symbol Error Rates for Linaer Block Codes," *IEEE Trans. Commun.*, vol. 36, pp. 613–617, May 1988.

15. J.-J. Chang, D.-J. Hwang, and M.-C. Lin, "Some Extended Results on the Search for Good Convolutional Codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1682–1697, September 1997.

16. D. Torrieri, *Principles of Secure Communication Systems*, 2nd ed. Boston: Artech House, 1992.

17. T. H. Liew and L. Hanzo, "Space-Time Codes and Concatenated Channel Codes for Wireless Communications," *Proc. IEEE*, vol. 90, pp. 187–219, February 2002.

18. S. W. Golomb, *Shift Register Sequences*, revised ed. Laguna Hills, CA: Aegean Park Press, 1982.

19. D. V. Sarwate and M. B. Pursley, "Crosscorrelation Properties of Pseudorandom and Related Sequences," *Proc. IEEE*, vol. 68, pp. 593–619, May 1980.

20. D. Torrieri, "Performance of Direct-Sequence Systems with Long Pseudonoise Sequences," *IEEE J. Select. Areas Commun.*, SAC-10, pp. 770–781, May 1992.

21. R. B. Ash and C. A. Doleans-Dade, *Probability and Measure Theory, 2nd ed.* San Diego: Academic Press, 2000.

22. C. C. Lee and R. Steele, "Closed-loop Power Control in CDMA Systems," *IEE Proc.-Commun.*, vol. 143, pp. 231–239, August 1996.

23. S. DeFina and P. Lombardo, "Error Probability Analysis for CDMA Systems with Closed-Loop Power Control," *IEEE Trans. Commun.*, vol. 49, pp. 1801–1811, October 2001.

24. D. Torrieri, "Instantaneous and Local-Mean Power Control for Direct-Sequence CDMA Cellular Networks," *IEEE Trans. Commun.*, vol. 50, pp. 1310–1315, August 2002.

25. M. Zorzi, "On the Analytical Computation of the Interference Statistics with Applications to the Performance Evaluation of Mobile Radio Systems," *IEEE Trans. Commun.*, vol. 45, pp. 103–109, January 1997.

26. A. J. Viterbi, *CDMA Principles of Spread Spectrum Communication*. Reading, MA: Addison-Wesley, 1995.

27. G. E. Corazza, G. DeMaio, and F. Vatalaro, "CDMA Cellular Systems Performance with Fading, Shadowing, and Imperfect Power Control," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 450–459, May 1998.

28. R. N. McDonough and A. D. Whalen, *Detection of Signals in Noise, 2nd ed.* San Diego: Academic Press, 1995.

29. S. Verdu, *Multiuser Detection*. New York: Cambridge University Press, 1998.

30. M. Sawahashi et al., "Experiments on Pilot Symbol-Assisted Coherent Multistage Interference Canceller for DS-CDMA Mobile Radio," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 433–449, February 2002.

31. L. B. Milstein, "Wideband Code Division Multiple Access," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1344–1354, August 2000.

32. D. Torrieri, "Mobile Frequency-Hopping CDMA Systems," *IEEE Trans. Commun.*, vol. 48, pp. 1318–1327, August 2000.

33. F. Adachi and J. D. Parsons, "Unified Analysis of Postdetection Diversity for Binary Digital FM Radio," *IEEE Trans. Veh. Technol.*, vol. 37, pp. 189–198, November 1988.

34. A. A. Shaar and P. A. Davies, "A Survey of One-Coincidence Sequences for Frequency Hopped Spread Spectrum Systems," *Proc. IEE, Pt. F, Commun., Radar Signal Processing*, pp. 719–724, December 1984.

35. S. Chennakeshu et al., "Capacity Analysis of a TDMA-Based Slow-Frequency-Hopped Cellular System," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 531–542, August 1996.

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2002	Final, Oct. 2001 to Sept. 2001

4. TITLE AND SUBTITLE
Spread-Spectrum Code-Division Multiple Access

5. FUNDING NUMBERS
DA PR: N/A
PE: 62120A

6. AUTHOR(S)
Don Torrieri

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
Attn: AMSRL- CI-C email: dtorrieri@arl.army.mil
2800 Powder Mill Road
Adelphi, MD 20783-1197

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-2721

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, MD 20783-1197

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
ARL PR: 2FOVF4
AMS code: 622120H1600

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report presents a detailed analysis of the most important aspects of spread-spectrum code-division multiple access (CDMA). The foundations are provided to enable the reader to understand the more advanced literature. The fundamental characteristics of a fading channel are summarized. Various types of diversity and diversity combining are examined. Both direct-sequence and frequency-hopping systems and networks are analyzed.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Fading, code-division multiple-access, CDMA, diversity	150
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL